



Causal Mediation Analysis With Time-Varying and Multiple Mediators

Citation

Lin, Sheng-Hsuan. 2016. Causal Mediation Analysis With Time-Varying and Multiple Mediators. Doctoral dissertation, Harvard T.H. Chan School of Public Health.

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:27201718>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

**CAUSAL MEDIATION ANALYSIS WITH TIME-VARYING AND MULTIPLE
MEDIATORS**

Sheng-Hsuan Lin

A Dissertation Submitted to the Faculty of
The Harvard T.H. Chan School of Public Health
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Science
in the Departments of Epidemiology
Harvard University
Boston, Massachusetts.

May, 2016

CAUSAL MEDIATION ANALYSIS WITH TIME-VARYING AND MULTIPLE MEDIATORS

Abstract

The assessment of direct and indirect effects with time-varying and multiple mediators is a common but challenging problem, and standard mediation analysis approaches are generally not applicable in this context. This dissertation focused on extending mediation analysis into a setting with time-varying and multiple mediators. An interventional approach has been used to define and identify the direct and indirect effects as well as path specific effects based in a causal inference framework, propose a parametric approach to estimate these effects, and provide an algorithm as well as corresponding software for practical application.

In the first paper, we develop a parametric estimation approach to the mediational g-formula, including a feasible algorithm implemented in a freely available SAS macro. In the Framingham Heart Study data, we apply this method to estimate the interventional analogues of natural direct and indirect effects of smoking behaviors sustained over a 10-year period on blood pressure when considering weight change as a time-varying mediator. Compared with non-smoking, smoking 20 cigarettes per day for 10 years was estimated to increase blood pressure by 1.18 (95 % CI: -0.68, 2.69) mm-Hg. The direct

effect was estimated to increase blood pressure by 1.52 (95 % CI: -0.25, 2.90) mm-Hg, and the indirect effect was -0.34 (95 % CI: -0.52, -0.13) mm-Hg, which is negative because smoking leads to lower weight which leads to lower blood pressure. These results provide evidence that weight change in fact partially conceals the detrimental effects of cigarette smoking on blood pressure. Our work represents the first application of the parametric mediational g-formula in an epidemiologic cohort study.

The second paper proposes an approach to conduct mediation analysis for survival data with time-varying exposures, mediators, and confounders. We identify the direct and indirect effects through a survival mediational g-formula and provide the required assumptions. We also provide a feasible parametric approach along with an algorithm and software to estimate these effects. We apply this method to analyze the Framingham Heart Study data to investigate the causal mechanism of smoking on mortality. The risk ratio of smoking 30 cigarettes per day for ten years compared with no smoking on mortality is 2.34 (95 % CI = (1.44, 3.70)). Of the effect, 7.91 % is mediated by coronary artery disease. The survival mediational g-formula constitutes a powerful tool for conducting mediation analysis with longitudinal data.

Finally, the third paper further proposes a method, defining a randomly interventional analogue of path-specific effect, which can always be non-parametrically identified under assumptions of no unmeasured confounding. This method also allows settings with mediators dependent on each other, interaction, and mediator-outcome confounders which are affected by exposure. In addition, under linearity and no-interaction, our method has the same form of traditional path analysis for path-specific effect. Furthermore, under a single mediator without a mediator-outcome confounder

affected by exposure, it also has the same form of the results of causal mediation analysis. We also provide SAS code for settings of linear regression with exposure-mediator interaction and perform analysis in Framingham Heart Study dataset, investigating the mechanism of smoking on systolic blood pressure mediated by both cholesterol and body weight. Allowing decomposition of total effect into several analogues of path-specific effects, our method contributes to the investigation of complicated causal mechanisms in settings with multiple mediators.

Key words: causal inference; effect decomposition; mediation analysis; path-specific effect; survival outcome; time-varying; multiple mediators; g-formula; parametric approach; stochastic intervention.

TABLE OF CONTENTS

Supplementary Files List.....	viii
ACKNOWLEDGEMENTS	ix
Paper 1 Parametric mediational g-formula approach to mediation analysis with time-varying exposures, mediators, and confounders: an application for smoking, weight, and blood pressure	1
Abstract	2
Introduction	3
<i>Case study for Framingham dataset: smoking, weight, and blood pressure</i>	4
Methods Development	5
<i>Notation and review for causal mediation analysis</i>	5
<i>Notation and review of mediation analysis with time-varying mediators and the mediational g-formula</i>	9
<i>Parametric mediational g-formula</i>	12
Data Application.....	16
Results	18
Discussion	19
<i>Conclusion</i>	22
Reference.....	22
Table 1.1. Summary of covariate models.	29
Table 1.2. Baseline characteristics of eligible participants grouped by former smoking status.	31
Table 1.3. Estimates of the total effect of smoking 20 cigarettes per day for 10 years (compared with no smoking) on SBP and BMI.	33
Table 1.4. Randomly interventional analogue of total effect, of natural direct effect, and of natural indirect effect for the effect of smoking 20 cigarettes per day for 10 years (compared with no smoking) on SBP, mediated by BMI change over time.	34
Figure 1.1.Simple model for mediation analysis.	35
Figure 1.2.Mediation analysis with a mediator-outcome confounder L that is affected by exposure.....	36
Figure 1.3. Time-varying mediation with ordering of variables of A(t), M(t), L(t), for t = 0 to T-1.	37
Paper 2 Mediation analysis for survival outcome with time-varying exposures, mediators, and confounders: a case study of the Framingham Heart Study	38
Abstract	39

Introduction	40
<i>Background</i>	40
<i>Case study for Framingham dataset: smoking, coronary artery disease, and overall mortality</i>	42
Notations and definitions.....	42
<i>Notations</i>	42
<i>Definitions of mediation parameter and the direct and indirect effects</i>	44
Identification and estimation of survival mediational g-formula	48
<i>Assumptions and identification for survival mediational g-formula</i>	48
<i>Parametric approach for survival mediational g-formula</i>	53
Analysis of Framingham Heart Study data	55
Discussion	58
<i>Conclusion</i>	59
Reference.....	60
Figure 2.1. Time-varying mediation with variables of $A(t)$, $M(t)$, $L(t)$, and $S(t)$, for $t = 1$ to T .63	
Figure 2.2. Simple model for survival mediation analysis with fixed exposure and time-varying mediators with 2 timepoints.	64
Figure 2.3. Simple model for survival mediation analysis with fixed exposure and time-varying mediators.	65
Table 2.1. Estimates of the overall effect of smoking 30 cigarettes per day for 10 year (compared with no smoking) on mortality.	66
Table 2.2. Mediation analysis for the effect of smoking 30 cigarettes per day for 10 years (compared with no smoking) on overall mortality, mediated by cardiovascular disease.....	67
Paper 3 Interventional Approach for Path-Specific Effects	68
Abstract	69
Introduction	70
Notation and review of standard causal mediation analysis in a setting with two mediators....	71
<i>Notation and review for counterfactual models</i>	71
<i>Definitions of total effect (TE), control direct effect (CDE), natural direct effect (NDE), natural indirect effect (NDE), and the randomly interventional analogues of these effects</i> .	72
Definition and identification of randomly interventional analogues of path-specific effects ...	77
<i>Relation to other causal mediation analysis studies</i>	82
A regression based approach and illustration	84

<i>Illustration</i>	86
Discussion	88
Reference.....	90
Figure 3.1. Causal diagrams for a setting with two mediators and time-varying mediator-outcome confounders affected by exposure.....	93
Figure 3.2. Causal diagrams for a setting with two mediators but no time-varying mediator-outcome confounders.	94
Table 3.1. Proportions of the effect of smoking on systolic blood pressure mediated by cholesterol and/or body mass index.	95
Supplementary Files	96

Supplementary Files List

Supplementary files of paper 1

Appendix 1: Description of mgformula SAS macro for parametric mediational g-formula

Supplementary files of paper 2

Appendix 2.1: Identification of mediation parameter to the survival mediational g-formula

Appendix 2.2: Identification for traditional definition of mediation parameter under the sequential monotonicity assumption

Appendix 2.3 : Stochastic process version

Appendix 2.4: Assumption 3 with unmeasured survival confounders

Supplementary files of paper 3

Appendix 3.1: Identification of randomly interventional analogue of mediation parameter

Appendix 3.2: A setting with three mediators

Appendix 3.3: A regression-based approach

Appendix 3.4: SAS code for randomly interventional analogues of path-specific effects

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. Tyler VanderWeele, the most perfect adviser I can imagine, for his great advice and support for my research and dissertation. I would also like to thank my two committee members: Dr. Eric Tchetgen Tchetgen, for his great training in my second year at Harvard, and Dr. Miguel Hernan, for providing the FHS data and his amazing teaching in my first two causal inference courses. I would like to thank Roger Logan, for always coming to the rescue whenever I needed SAS debug; Jessica and Priyanka, for their suggestions; and Kathy Brenner, for her excellent help in academic writing.

I would like to thank Dr. Jung-Der Wang, my adviser in Taiwan, for introducing me to the field of Public health and Epidemiology, and for being a role model both academic-wise and life-wise. I would like to thank Dr. Yen-Tsung Huang for his guidance and advice for my academic career in methodology.

I would like to thank all my friends in Boston for being kind to me. It means a lot to me and has made the past 5 years an enjoyable journey. I especially thank Dr. Chen-Hao Chen for the inspiring discussions; Dr. Chen-Yu Wang and Dr. Chuan-Chin Huang for their critical help during my first year at Harvard. In addition, I would like to thank those who lighted me up when I was drowning in the dark. These wonderful people are Bao-Lao, Yu-Han, Yi-Han, Yu-Wei, Yen-Feng, Chang-Zheng, and Peng Cheng, etc. I will always remember my great classmates, Ding Ming and MingYang, etc., and the great time we had preparing for the qualify exam and discussing homework together.

I would like to thank my parents, Hsin-Tsai Lin and Yueh-Jen Liao, for raising me up, always believing in me, and being supportive for my decisions throughout the critical stages of my life. Many thanks to my grandparents, my brothers, and all my family, for their big support. I would also like to thank my wife, Shih-Chen, for always keeping me company. Without her, I would just end up being another ordinary clinician in Taiwan and would never have enough courage to study abroad. I am so blessed to have her in my life.

Lastly, I would like to thank my beautiful homeland, Taiwan, the place I grew up in, and also the government for providing me the four-year financial support. I promise in the near future, I will return to her embrace and contribute the rest of my life to make her the best country in the world to live in.

**Paper 1 Parametric mediational g-formula approach to mediation analysis with
time-varying exposures, mediators, and confounders: an application for smoking,
weight, and blood pressure**

Sheng-Hsuan Lin^{1,2,*}, Jessica Young¹, Roger Logan¹, Eric J. Tchetgen Tchetgen^{1,2}, Tyler
J. VanderWeele^{1,2}

*: corresponding author

¹Department of Epidemiology, Harvard School of Public Health, Boston, MA

²Department of Biostatistics, Harvard School of Public Health, Boston, MA

Abstract

The assessment of direct and indirect effects with time-varying mediators and confounders is a common but challenging problem, and standard mediation analysis approaches are generally not applicable in this context. The mediational g-formula was recently proposed to address this problem, paired with a semi-parametric estimation approach to evaluate longitudinal mediation effects empirically. In this paper, we develop a parametric estimation approach to the mediational g-formula, including a feasible algorithm implemented in a freely available SAS macro. In the Framingham Heart Study data, we apply this method to estimate the interventional analogues of natural direct and indirect effects of smoking behaviors sustained over a 10-year period on blood pressure when considering weight change as a time-varying mediator. Compared with non-smoking, smoking 20 cigarettes per day for 10 years was estimated to increase blood pressure by 1.18 (95 % CI: -0.68, 2.69) mm-Hg. The direct effect was estimated to increase blood pressure by 1.52 (95 % CI: -0.25, 2.90) mm-Hg, and the indirect effect was -0.34 (95 % CI: -0.52, -0.13) mm-Hg, which is negative because smoking leads to lower weight which leads to lower blood pressure. These results provide evidence that weight change in fact partially conceals the detrimental effects of cigarette smoking on blood pressure. Our work represents the first application of the parametric mediational g-formula in an epidemiologic cohort study.

Introduction

Mediation analysis, a method to decompose the total effect of an exposure on an outcome into direct and indirect effects through a mediator, is essential for investigating pathways or mechanisms in epidemiology and in the social sciences. Causal mediation analysis, defining natural direct and indirect effects based on counterfactual models, extends traditional mediation analysis to settings involving nonlinearities and interactions^{1,2}. Numerous methodological approaches based on causal mediation analysis have been developed in recent years allowing different outcome scales, including additive, multiplicative, and odds ratio scales, as well as other models for time to event data³⁻⁸. Most of the approaches mentioned above only consider a point exposure and a subsequent point mediator. When conducting causal mediation analysis with longitudinal data, the restriction of only one single exposure and mediator neglects exposures or mediators at other time-points, which thus potentially results in loss of valuable information.

Robins has proposed the g-formula to estimate the total effect in settings with time-varying exposures and confounders⁹. In addition, when mediators can be intervened upon, the g-formula can estimate the controlled direct effect (CDE) by comparing the effects of two different exposure levels and specifying the mediators at certain fixed values. For mediation, however, the natural direct effect (NDE) and the natural indirect effect (NIE) involved in effect decomposition are, unfortunately, not identified. VanderWeele and Tchetgen Tchetgen¹⁰ proposed the mediational g-formula to overcome the methodological challenges of causal mediation analysis with time-varying mediators. As discussed below, this method decomposes the randomized interventional analogue of

total effect (rTE) into interventional analogues of natural direct effect (rNDE) and indirect effect (rNIE) and moves beyond the limitations of a single exposure and a single mediator. Time-varying confounders can also be accounted for provided no unobserved confounding is present. For estimation of the rNDE and rNIE, VanderWeele and Tchetgen Tchetgen proposed a semi-parametric approach based on inverse probability weighted (IPW) estimators. As with standard IPW, this approach is potentially unstable if the exposure is continuous or the weights are highly variable. Instead, we consider an alternative approach that is potentially more stable and efficient, by implementing a fully parametric mediational g-formula approach, using a user-friendly algorithm implemented in freely available software. We then apply this method to the Framingham Heart Study (FHS) data to investigate the effect of smoking on blood pressure mediated by weight change.

Case study for Framingham dataset: smoking, weight, and blood pressure

In past research, the association between smoking and blood pressure has been controversial¹¹⁻¹⁶. In some studies, average blood pressure, as measured using a domestic blood pressure monitor¹⁷⁻¹⁹, is lower among smokers than among non-smokers at particular times of the day^{14,20}. According to most literature, among former smokers, smoking cessation increases blood pressure^{21,22} while some studies fail to find this association significant²³.

Three possible mechanisms might explain this association. First, smoking activates the autonomic nervous system (ANS) to increase blood pressure directly, but smoking

cessation also activates the ANS, increasing blood pressure among former smokers^{24,25}. Second, smoking elevates blood pressure through exacerbating arterial stiffness. Third, smoking decreases blood pressure through weight loss²¹. The adverse effect of smoking on elevated blood pressure might thus be partially concealed by weight loss. Because both smoking status and weight change vary over time, the mediational g-formula is an can be used to study the extent of this adverse effect mediated by weight loss. In this study, we obtain the estimates by applying our method to the FHS data and demonstrate an application of causal mediation analysis with time-varying mediators.

Methods Development

Notation and review for causal mediation analysis

First consider a setting with an exposure, mediator, and outcome measured at a single time. Let A denote an exposure, Y an outcome, and M a mediator. Let V denote a set of baseline covariates not affected by the exposure. The relations among these variables are described in Figure 1.1. Under counterfactual models^{26,27}, Y_a and M_a denote the counterfactual values of the outcome and the mediator, respectively, if exposure A is set to level a . Y_{am} denotes the counterfactual value of the outcome if exposure A is set to level a , and mediator M is set to level m . In addition, $Y_{aM_a^*}$ denotes the counterfactual value of the outcome if exposure A is set to level a and mediator M is set to level M_a^* . Under the consistency assumption^{6,28,29}, $Y_a = Y$ and $M_a = M$ if $A = a$; $Y_{am} = Y$ if $A = a$ and $M = m$; and the composition assumption that $Y_{aM_a^*} = Y$ if $A = a$ and $M = M_a^*$.

Mediation analysis decomposes the overall effect into a direct effect (the effect not through the mediator) and an indirect effect (the effect through the mediator). Under the above counterfactual models, causal mediation analysis usually defines the TE, NDE, and NIE to represent the overall, direct, and indirect effects, respectively. Let $A = a$ and $A = a^*$ denote two hypothetical intervention statuses, exposure and non-exposure, respectively. The TE is defined as $E[Y_a - Y_{a^*}]$ or equivalently, $E[Y_{aMa} - Y_{a^*Ma^*}]$; the NDE is defined as $E[Y_{aMa^*} - Y_{a^*Ma^*}]$, and the NIE is defined as $E[Y_{aMa} - Y_{aMa^*}]$. These effects have been described extensively elsewhere and arguments have been made that they are theoretically appealing for effect decomposition^{1,2,6}. However, in the presence of time-varying confounders, the NDE and NIE are not generally identified from data even if these confounders are observed^{30,31}. To address the problem of identification, an alternative definition uses the rTE, rNDE, and rNIE to represent the overall, direct, and indirect effects, respectively³²⁻³⁴. Let G_a denote a random draw from the distribution of the mediator amongst those with exposure status $A = a$. When exposure is set to a (or a^*), the distribution of mediator among whole population is also determined. Therefore, for every individual, a random draw from this distribution, G_a (or G_{a^*}), will be independent of the counterfactual mediator, M_a (or M_{a^*}) and outcome (Y_{am}). The rTE is defined as $E[Y_{aG_a}] - E[Y_{a^*G_{a^*}}]$. The rNDE is defined as $E[Y_{aG_a}] - E[Y_{a^*G_{a^*}}]$, which can be interpreted as an effect on the outcome comparing exposure versus no exposure with the mediator in both cases randomly drawn from the distribution of the population when given no exposure. Finally, the rNIE is defined as $E[Y_{aG_a}] - E[Y_{aG_{a^*}}]$, interpreted as an effect on the outcome of randomly assigning an individual who is given exposure to a value of the mediator drawn from the distribution of the mediator amongst those given

exposure versus not given exposure. We have the decomposition: $E[Y_{aGa}] - E[Y_{a^*Ga^*}] = (E[Y_{aGa}] - E[Y_{aGa^*}]) + (E[Y_{aGa^*}] - E[Y_{a^*Ga^*}])$, so the overall effect decomposes into the sum of the effect through the mediator (i.e. the indirect effect) and the direct effect.

For the NDE and the NIE, both effects can be identified under four assumptions: (1) $Y_{am} \perp A|V$ (no unmeasured exposure-outcome confounder); (2) $Y_{am} \perp M|V, A$ (no unmeasured mediator-outcome confounder); (3) $M_a \perp A|V$ (no unmeasured exposure-mediator confounder); and (4) $Y_{am} \perp M_{a^*}|V$ (no mediator-outcome confounder affected by exposure)^{6,30}, when \perp denotes independence and $X \perp Y | Z$ denotes that X is independent of Y conditional on Z . Under these assumptions, NDE and NIE are identified by the following expressions:

$$\begin{aligned} \text{NDE} = & \\ \sum_v \sum_m \{E[Y|a, m, v] - E[Y|a^*, m, v]\} P(m|a^*, v) P(v) & \end{aligned} \quad (1)$$

$$\begin{aligned} \text{NIE} = & \\ \sum_v \sum_m E[Y|a, m, v] \{P(m|a, v) - P(m|a^*, v)\} P(v) & \end{aligned} \quad (2)$$

The fourth assumption holds only under a non-parametric structural equation model and would be violated under several settings. The most common one is the existence of a mediator-outcome confounder L that is affected by exposure (Figure 1.2), in which case

the fourth assumption fails, even if this confounder is observed. When this confounder is a single binary variable, Tchetgen Tchetgen and VanderWeele proposed a method to identify the NDE and NIE by assuming the monotonicity about the effect of exposure on this confounder ³⁵. A severe shortcoming of this assumption is that even if the mediator is restricted to occurring immediately after the exposure, the assumption cannot be ensured. As a result, this approach is not generally applicable for settings with time-varying mediators. However, even if this assumption fails, given assumptions (1) to (3) hold and the mediator-outcome confounder is observed, the rNDE and rNIE in the second definition are still identifiable from the data by the empirical expressions given by ³⁴:

$$\begin{aligned} \text{rNDE} = & \\ \sum_{v,l,m} \{ & E[Y|a, l, m, v]P(l|a, v) - E[Y|a^*, l, m, v]P(l|a^*, v) \} P(m|a^*, v)P(v) \end{aligned} \quad (3)$$

$$\begin{aligned} \text{rNIE} = & \\ \sum_{v,l,m} E[Y|a, l, m, v]P(l|a, v) \{ & P(m|a, v) - P(m|a^*, v) \} P(v) \end{aligned} \quad (4)$$

To understand the difference in the effects better, note that the NDE and NIE cannot be checked by designing a randomized trial even if we were able to intervene on both the exposure and the mediator. However, it is possible to design a two-stage trial to estimate rNDE and rNIE. A randomized trial could first be conducted to obtain the empirical distribution of counterfactual mediator given exposed and non-exposed by randomizing

the exposure and measuring the mediator distributions. We can then estimate rNIE in a second trial by the effect on the outcome of assigning an individual who is given the exposure to a value of the mediator sampled from the marginal distribution of the mediator amongst those given exposure versus no exposure, using the empirical distributions of the mediator estimated from the first randomized trial. Similarly, we can estimate rNDE by a direct effect comparing exposure versus no exposure with the mediator in both cases randomly drawn from the empirical distribution of the population when given no exposure, which was also estimated from the first randomized trial.

Notation and review of mediation analysis with time-varying mediators and the mediational g-formula

Consider exposures, mediators, and confounders that vary over time in longitudinal settings with T measurements at time $t = 0, 1, 2, \dots, T-1$. Let $(A(0), A(1), \dots, A(T-1))$, $(M(0), M(1), \dots, M(T-1))$, and $(L(0), \dots, L(T-1))$ denote values of the time-varying exposures, mediators, and confounders at periods $1, \dots, T$, with the final outcome of interest Y . The initial baseline confounders are included in $L(0)$. Figure 1.3 depicts a possible data generating mechanism under which these assumptions would hold.

For any variable W , let $\overline{W(t)} = (W(0), W(1), \dots, W(t))$ and let $\overline{W} = \overline{W(T-1)} = (W(0), W(1), \dots, W(T-1))$. Let $Y_{\overline{a}\overline{m}}$ be the counterfactual value of Y given \overline{A} is set to \overline{a} and \overline{M} is set to \overline{m} . Let $M_{\overline{a}}(t)$ be the counterfactual value of $M(t)$ given $\overline{A(t)}$ is set to $\overline{a(t)}$. Let $G_{\overline{a}}(t)$ denote a random draw from the distribution of the mediator $M_{\overline{a}}(t)$. Let $\overline{A} = \overline{a}$ and $\overline{A} = \overline{a^*}$ denote two hypothetical intervention statuses, for example, exposed

from $t = 0$ to $T-1$ and non-exposed from $t = 0$ to $T-1$, respectively. In this setting, we define TE, NDE, and NIE as $E[Y_{\bar{a}}] - E[Y_{\bar{a}^*}]$ (i.e. $E[Y_{\bar{a}\bar{M}\bar{a}}] - E[Y_{\bar{a}^*\bar{M}\bar{a}^*}]$), $E[Y_{\bar{a}\bar{M}\bar{a}}] - E[Y_{\bar{a}^*\bar{M}\bar{a}^*}]$, and $E[Y_{\bar{a}\bar{M}\bar{a}}] - E[Y_{\bar{a}\bar{M}\bar{a}^*}]$, respectively; while rTE, rNDE, and rNIE are defined as $E[Y_{\bar{a}\bar{G}\bar{a}}] - E[Y_{\bar{a}^*\bar{G}\bar{a}^*}]$, $E[Y_{\bar{a}\bar{G}\bar{a}^*}] - E[Y_{\bar{a}^*\bar{G}\bar{a}^*}]$, and $E[Y_{\bar{a}\bar{G}\bar{a}}] - E[Y_{\bar{a}\bar{G}\bar{a}^*}]$, respectively. We also can decompose the TE into the NDE and NIE. Similarly, rTE is decomposed into rNDE and rNIE (i.e., $E[Y_{\bar{a}\bar{G}\bar{a}}] - E[Y_{\bar{a}^*\bar{G}\bar{a}^*}] = E[Y_{\bar{a}\bar{G}\bar{a}^*}] - E[Y_{\bar{a}^*\bar{G}\bar{a}^*}] + E[Y_{\bar{a}\bar{G}\bar{a}}] - E[Y_{\bar{a}\bar{G}\bar{a}^*}]$).

If the entire vector $\bar{A} = \overline{A(T-1)}$ is taken as the joint exposure of interest and $\bar{M} = \overline{M(T-1)}$ as the mediators of interest, then assumption (4) is violated because the variable $L(1)$ is affected by $A(0)$ and confounds the mediator-outcome relationship between $M(1)$ and Y (similarly, the $L(t)$ is affected by $\overline{A(t-1)}$ and confounds the relationship between $M(t)$ and Y , when $t = 1, \dots, T-1$). Therefore, NDE and NIE cannot be identified in this setting. However, rNDE and rNIE are still identifiable under the following three assumptions for all t : (1) $Y_{\bar{a}\bar{m}} \perp A(t) | \overline{A(t-1)}, \overline{M(t-1)}, \overline{L(t)}$ (no exposure-outcome confounding conditional on the past variables), (2) $Y_{\bar{a}\bar{m}} \perp M(t) | \overline{A(t)}, \overline{M(t-1)}, \overline{L(t)}$ (no mediator-outcome confounding conditional on the past variables), and (3) $M_{\bar{a}}(t) \perp A(t) | \overline{A(t-1)}, \overline{M(t-1)}, \overline{L(t)}$ (no exposure-mediator confounding conditional on the past variables)¹⁰. Given the three assumptions, VanderWeele and Tchetgen Tchetgen¹⁰ show that the rNDE and rNIE are identified non-parametrically by the following equations:

$$rNDE = Q(\bar{a}, \bar{a}^*) - Q(\bar{a}^*, \bar{a}^*)$$

(5)

$$rNIE = Q(\bar{a}, \bar{a}) - Q(\bar{a}, \bar{a}^*)$$

(6)

where $Q(\bar{a}_1, \bar{a}_2)$

=

$$\begin{aligned} & \sum_{\bar{m}} \sum_{\bar{l}} E[Y | \bar{a}_1, \bar{m}, \bar{l}] \prod_{t=0}^{T-1} P(l(t) | \bar{a}_1(t-1), \bar{m}(t-1), \bar{l}(t-1)) \\ & \times \sum_{\bar{l}'} \prod_{t=0}^{T-1} P(m(t) | \bar{a}_2(t), \bar{m}(t-1), \bar{l}'(t)) P(l'(t) | \bar{a}_2(t-1), \bar{m}(t-1), \bar{l}(t-1)) \end{aligned} \quad (7)$$

As in VanderWeele and Tchetgen Tchetgen, we refer to this expression $Q(\bar{a}_1, \bar{a}_2)$ above as the mediational g-formula. When there are no mediators, i.e. \bar{M} equals empty, this formula (7) reduces to the standard g-formula:

$$Q(\bar{a}) =$$

$$\sum_{\bar{l}} E[Y | \bar{a}_1, \bar{l}] \prod_{t=0}^{T-1} P(l(t) | \bar{a}_1(t-1), \bar{l}(t-1)) \quad (8)$$

Further, rNDE (5) and rNIE (6) reduce to the NDE and NIE, respectively, when there are no time-varying confounders¹⁰.

Parametric mediational g-formula

VanderWeele and Tchetgen Tchetgen described how to use IPW of marginal structural models (MSM) to estimate the mediational g-formula $Q(\bar{a}_1, \bar{a}_2)$ (7) in realistic high-dimensional settings¹⁰. However, this approach can perform poorly with continuous exposures and mediators and can also be inefficient. As an alternative, an adaptation of the standard parametric g-formula^{9,36} can be used to parametrically estimate the mediational g-formula in high-dimensional settings and, in turn, the rNDE (5) and rNIE (6).

We begin by briefly reviewing the standard parametric g-formula. This approach parametrically estimates the standard g-formula by (i) fitting parametric models for the joint density of the outcome and time-varying covariates and (ii) using the estimated parameters of these models to simulate many covariate histories consistent with the exposure intervention \bar{a} . Specifically, the following algorithm, which is implemented in a publicly available SAS macro³⁷, can be used to parametrically estimate the standard g-formula $Q(\bar{a})$:

(1) Fit parametric models for the observed data:

(1a) For times $t \geq 0$, fit parametric models for the joint density of the confounders and exposures at t given the measured past.

(1b) Fit a parametric model for the mean of the outcome at the end of follow-up given the measured past.

(2) Set baseline confounders and exposures to the observed sample values. Recursively, for each subject $i = 1, \dots, n$ and for each time $t = 0, 1, 2, \dots, T-1$:

- (2a) For $t \geq 0$, generate time t confounders and exposures based on the estimated model coefficients of (1a) and previously generated exposures and confounders under intervention.
- (2b) Assign time t exposures under intervention \bar{a} .
- (3) Simulate the outcome for each of the n generated histories in step 2 based on the estimated model coefficients of (1b).
- (4) Take the mean over n simulated outcomes in (3) to estimate $Q(\bar{a})$.

Here, we have adapted the above algorithm and associated SAS macro code to parametrically estimate the mediational g-formula $Q(\bar{a}_1, \bar{a}_2)$, the rTE, the rNDE, and the rNIE. The primary difference between the parametric g-formula and the parametric mediational g-formula is, under the latter algorithm, the estimated model coefficients from step (1) are additionally used to estimate the joint distribution of the time-varying mediators (marginal over all other covariates) under both exposure interventions \bar{a} and \bar{a}^* . These are then used to assign values of the mediator under the joint exposure and mediator interventions $(\bar{a}, \bar{G}_{\bar{a}})$, $(\bar{a}, \bar{G}_{\bar{a}^*})$, $(\bar{a}^*, \bar{G}_{\bar{a}})$ and $(\bar{a}^*, \bar{G}_{\bar{a}^*})$. The algorithm is as follows:

- (1) Fit parametric models for the observed data:

- (1a) For times $t \geq 0$, fit parametric models for the joint density of the confounders, exposures and mediators at t given the measured past.

- (1b) Fit a parametric model for the mean of the outcome at the end of follow-up given the measured past.
- (2) Estimate the joint distribution of time-varying mediators under time-varying exposure interventions \bar{a} and \bar{a}^* :
- (2a) Set baseline ($t = 0$) covariates to the observed values for subject i . Recursively, for each time $t = 0, \dots, T-1$ and each subject $i = 1, \dots, n$:
- (2a.i) For $t \geq 0$, generate time t confounders, exposure, and mediator based on the estimated model coefficients of (1a) and previously generated covariates under the time-varying exposure intervention \bar{a} through $t-1$.
- (2a.ii) Assign time t exposure under the intervention \bar{a} .
- (2b) For each $t = 0, \dots, T-1$, randomly permute the n values of the joint mediators assigned under intervention \bar{a} in (2a). For each t , save this permutation for use in (3) below (we obtain $\bar{G}_{\bar{a}}$ in this step).
- (2c) Repeat (2a) replacing intervention \bar{a} with \bar{a}^* .
- (2d) Repeat (2b) replacing intervention \bar{a} with \bar{a}^* (we obtain $\bar{G}_{\bar{a}^*}$ in this step).
- (3) Estimate $Q(\bar{a}, \bar{a})$, $Q(\bar{a}, \bar{a}^*)$, $Q(\bar{a}^*, \bar{a})$ and $Q(\bar{a}^*, \bar{a}^*)$ by repeating the following for each $(\bar{a}_1, \bar{a}_2) = (\bar{a}, \bar{a}), (\bar{a}, \bar{a}^*), (\bar{a}^*, \bar{a})$ and (\bar{a}^*, \bar{a}^*) :
- (3a) Recursively for each time $t = 0, \dots, T-1$ and each subject $i = 1, \dots, n$:

- (3a.i) Repeat (2a.i) but replacing “time-varying exposure intervention \bar{a} through $t-1$ ” with the joint “time-varying exposure and mediator intervention $(\bar{a}_1, \bar{G}_{\bar{a}_2})$ ”.
- (3a.ii) Assign the time t mediator as the i^{th} component of the permuted vector for time t from (2b) (if $\bar{a}_2 = \bar{a}$) or (2d) (if $\bar{a}_2 = \bar{a}^*$).
- (3a.iii) Assign time t exposure under the intervention \bar{a}_1 .
- (3b) Simulate the outcome given each of the $i = 1, \dots, n$ histories based on the estimated model coefficients of (1b) and the histories generated in (3a).
- (3c) Estimate $Q'(\bar{a}_1, \bar{a}_2)$ as the mean over the n simulated outcomes in (3b).
- (3d) Repeat (1) to (3c) for some fixed number K (e.g. 25) times, using different permutation in (2b) for each time.
- (3e) Estimate $Q(\bar{a}_1, \bar{a}_2)$ as the mean of the K (e.g. 25) values of $Q'(\bar{a}_1, \bar{a}_2)$ in (3d).

The algorithm can stop at (3c) and use the $Q'(\bar{a}_1, \bar{a}_2)$ as the unbiased estimate of $Q(\bar{a}_1, \bar{a}_2)$. However, the repeated steps in (3d) can improve standard errors for smaller sample sizes. Estimates of the rNDE and rNIE are then calculated from the estimates of the four $Q(\bar{a}_1, \bar{a}_2)$ in (3). 95% confidence intervals are calculated based on repeating the above algorithm in 500 bootstrap samples of the original n observations. This algorithm can be implemented with the `mgformula` macro, freely accessible with documentation at <http://www.hsph.harvard.edu/causal/software/>. Please see the Appendix 1 for details.

Data Application

Beginning in 1948 in Framingham, Massachusetts, the FHS is a longitudinal cohort study. The original cohort consists of 5,209 participants aged from 30 to 62 years old without cardiovascular disease (CVD) history at baseline. All the participants underwent examinations at the beginning of the study and routinely every two years after that. During each exam, potential CVD risk factors were collected, including socio-demographic data, lifestyle characteristics, detailed medical history, physical examination data, and blood samples. Further details on the design of FHS have been described elsewhere^{23,38}. The purpose of the analysis here is to illustrate the parametric meditational g-formula approach and software.

We specify exam 3 as the first exam and exams 1 and 2 as pre-baseline covariates to allow the function of the past in the models of step (1) of the estimation algorithm to depend on two lagged periods of the covariates. We follow the cohort for ten years (i.e. five visits) to reduce the proportion of death or loss to follow up to limit selection bias by death. Four exclusion criteria are listed below: (1) death or loss to follow up during the period before exam 7 (the end of follow-up); (2) no record at baseline on weight, height, smoking status, former smoking history, systolic blood pressure (SBP), or total cholesterol; (3) diagnosis of diabetes, cancer, or CVD at baseline; and (4) missing value for smoking status or BMI missing more than once. After these exclusions, 3,116 participants remain eligible for our analysis. For simplicity, we now refer to the original FHS exams 3,..., 7 as exams 1,..., 5.

SBP at exam 5 is the outcome Y . BMI during follow-up (exam 1 to 5) is the mediator \bar{M} . The exposure \bar{A} is smoking status during follow up, measured as

self-reported average number of cigarettes smoked per day. For missing BMI value or smoking status at a single time period, we carry forward the last observed value/status for one exam period only. We consider "smoking 20 cigarettes per day" and "no smoking" during follow-up as two hypothetical intervention levels $\bar{A} = \bar{a}$ and $\bar{A} = \bar{a}^*$. Time-varying covariates \bar{L} include the exam number, the systolic blood pressure (mm-Hg) at last exam, total non-fasting cholesterol level (mg/dl), and the usage of antihypertensive drugs. Baseline covariates include gender, age (years), height (meters), education ($\leq 8^{\text{th}}$ grade, some high school, high school graduate, some college, college graduate, post-graduate), occupation before retirement (executive/supervisory, technical, laborer, clerical, sales, housewife), marital status (single, married, widowed/divorced), and tobacco use at baseline (never user, current user, and past user). All covariates and the corresponding models are listed in Table 1.1.

The parametric g-formula is used to estimate the TE of smoking 20 cigarettes per day v.s. no smoking on SBP and on BMI at exam 5, by the g-formula macro. The parametric mediational g-formula is applied to conduct mediation analysis with time-varying mediators and exposures by our newly developed mgformula macro. We specify models for the outcome mean, as well as for each time-varying covariate (including mediator, exposure, and confounders) at each time point. We use current covariates and covariates at one period back (one lagged model) as the predictors. Specifically, we regress Y on main effects for $A(5)$, $M(5)$, $L(5)$, $A(4)$, $M(4)$, and $L(4)$. For $t = 0, 1, 2, \dots, 5$, we regress $L(t)$ on $A(t-1)$, $M(t-1)$, and $L(t-1)$; regress $M(t)$ on $A(t)$, $A(t-1)$, $M(t-1)$, $L(t)$, and $L(t-1)$; and regress $A(t)$ on $A(t-1)$, $M(t-1)$, $L(t)$, and $L(t-1)$.

(please refer to the Appendix 1 for more details). All analyses are conducted using SAS 9.4 (Cary, NC).

Results

The demographic and baseline health characteristics for smokers ($n = 1,759$), non-smokers ($n = 1,174$), and quitters ($n = 183$) are shown in Table 1.2. Compared with non-smokers, smokers have higher male proportion, younger age, and better education level. The majority of non-smokers are female (84.0%) and half of them have the occupation of housewife. The occupations of the smokers are mainly supervisor, laborer, and housewife. At baseline, the smokers appear to have better health status for lower SBP, cholesterol level, and BMI.

We use the g-formula to estimate the TE of smoking (20 cigarettes per day v.s. no smoking for 10 years) on SBP and BMI in Table 1.3. Smoking elevates SBP by 1.18 mm-Hg and reduces BMI by 0.16 kg/m^2 . In Table 1.4, we use the parametric mediational g-formula to simulate the SBP at the end of 10-year follow-up under no smoking with BMI distributed as the BMI under no smoking, smoking 20 cigarettes per day with BMI distributed as the BMI under smoking 20 cigarettes per day, no smoking with BMI distributed as the BMI under smoking 20 cigarettes per day, and smoking 20 cigarettes per day with BMI distributed as the BMI under no smoking. We then estimate the rTE, rNDE, and rNIE of smoking on SBP mediated by BMI. The rTE, rNDE, and rNIE estimates are 1.18 (95 % CI: -0.68, 2.69) mm-Hg, 1.52 (95 % CI: -0.25, 2.90) mm-Hg, and -0.34 (95 % CI: -0.52, -0.13) mm-Hg, respectively. The directions of rNIE and rNDE

are different, suggesting that the seemingly protective mediated effect of smoking through BMI may partially mask the detrimental direct effect of smoking on SBP.

Discussion

This is the first paper to provide a fully parametric method for causal mediation analysis with time-varying mediators. We develop an algorithm and corresponding SAS macro to use the mediational g-formula parametrically. We use the parametric approach to obtain estimates for the mediational g-formula by adapting the g-formula macro externally to build our SAS macro. Similar to the g-formula, we use Monte-Carlo simulation and bootstrapping methods for point and interval estimation, respectively. Since the estimation is an approximation of the maximum likelihood estimation, this estimation is asymptotically efficient provided the regression models are all correctly specified, while the IPW does not achieve the efficiency bound in a model where parametric assumptions about the weights and the MSM are correct. In addition, like other simulation-based methods³⁹, the approach here has the advantage of allowing for very flexible models such as quadratic linear models.

The parametric mediational g-formula provides a powerful method to investigate the mechanisms of an effect through time-varying mediators. Traditional techniques, allowing only one observation of the mediator and restricting it to variables occurring immediately after the exposure, inadequately capture the indirect effect when the exposures affect the mediators over time. One alternative approach is to estimate the controlled direct effect (CDE), the effect of the exposures on the outcome when fixing

the mediators at certain values. Obtained by applying the g-formula and specifying the mediators to these values, the CDE estimate provides valuable information for policy makers. For example, from FHS data we can provide the effect of smoking cessation decreasing SBP if BMI is fixed to a certain level. This is also described elaborately elsewhere⁴⁰. The effects from the mediational g-formula will, however, further allow for effect decomposition.

This study provides direct evidence for the hypothesis that “the adverse effect of smoking on high blood pressure is partially concealed by weight loss”^{13,21}. The concealment of the effect is partial because smoking does not decrease weight substantially (Table 1.3). Some studies report an effect of smoking cessation on increased blood pressure^{21,22}. The discrepancy with these studies might be attributed to a different analysis approach or simply the special characteristics of FHS participants.

Several limitations of this study should be noted. First, the use of rTE, rNDE, and rNIE, defined based on the stochastic interventions, results in different interpretations from the TE, NDE, and NIE³⁴. This is inevitable for causal mediation analysis with time-varying mediators because the NDE and NIE are not generally identified by empirical data in the presence of time-varying confounders, rNDE and rNIE can still serve as an analogue of NDE and NIE, and have the advantage that they can be verified by randomized controlled trial, while NDE and NIE cannot. Second, in the macro we have developed, the outcome can only be affected by the covariates at the most recent three exams because of a similar restriction of the g-formula macro that is employed for continuous/binary outcome types. This is not a limitation of the methodology itself, but only of the current implementation. The earlier covariates affect the outcome only

through these recent covariates. We specify the outcome variables in previous exams (i.e. the previous SBP values) as the time-varying covariates. Thus, the earlier covariates can also affect the outcome through the previous outcome variables. Third, selection bias has not been adjusted for here. Selection bias or truncation by death is a difficult problem for causal inference generally and for mediation analysis, even with only one single mediator. Most of the literature makes a sequential ignorability assumption that survival is effectively randomized conditional on the past. Under this assumption, the result can be interpreted as what would happen to the population if one could intervene to prevent death for everyone. Alternatively, one could also pursue sensitivity analysis approaches and there are also alternative stronger assumptions that would allow one to interpret the results without necessarily intervening upon death ^{41,42}. We hope to address this in future work but the extension of this to time-varying exposures and mediators is substantial. For the purposes of the illustration, we have simply focused on complete-case data and restricted our follow-up to ten years to partially address this selection bias. The relatively low proportion of loss to follow-up (< 20%) perhaps partially mitigates this problem. A general disadvantage of our proposed approach is that it may be particularly prone to bias due to model misspecification. Some misspecification may be theoretically guaranteed in complex longitudinal settings as considered here when the null hypothesis of no causal time-varying treatment effect is true, a problem known as the g-null paradox ⁴³. The magnitude of bias implied by the g-null paradox is not guaranteed and, depending on the setting, may be large or small. For some further consideration of the g-null paradox using numerical examples, see Young and Tchetgen Tchetgen ⁴⁴. The aforementioned IPW estimator offers an alternative method which is not subject to the g-null paradox and

might be used for the population estimands considered here⁴⁵. Finally, the analysis is subject to potential violation of the confounding assumptions. Future research could develop sensitivity analysis techniques for violations of these assumptions for our method.

Conclusion

The parametric mediational g-formula serves as a powerful and useful tool for mediation analysis with longitudinal data. Researchers can apply our method to disentangle the complicated causal mechanisms arising from time-varying mediators, exposures, and confounders. Further issues concerning the interpretation of the interventional direct and indirect effects can be found in VanderWeele and Tchetgen Tchetgen¹⁰. Using this method, we provide evidence that weight change in fact partially conceals the detrimental effect of cigarette smoking on blood pressure.

Reference

1. Pearl J. Direct and indirect effects. Proceedings of the seventeenth conference on uncertainty in artificial intelligence: Morgan Kaufmann Publishers Inc., 2001;411-420.
2. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 1992:143-155.
3. Tchetgen EJT, Shpitser I. Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis. *The Annals of*

- Statistics* 2012;**40**(3):1816-1845.
4. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological methods* 2013;**18**(2):137.
 5. van der Laan MJ, Petersen ML. Direct effect models. *The international journal of biostatistics* 2008;**4**(1).
 6. VanderWeele T, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface* 2009;**2**:457-468.
 7. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology (Cambridge, Mass.)* 2010;**21**(4):540.
 8. VanderWeele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *American journal of epidemiology* 2010;**172**(12):1339-1348.
 9. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling* 1986;**7**(9):1393-1512.
 10. VanderWeele TJ, Tchetgen Tchetgen E. Mediation Analysis with Time-Varying Exposures and Mediators. 2014.
 11. Chu N-F, Ding Y-A, Wang D-J, Shieh S-M. Relationship between smoking status and cardiovascular disease risk factors in young adult males in Taiwan. *Journal of cardiovascular Risk* 1996;**3**(2):205-208.
 12. Gordon T, Kannel WB, Dawber TR, McGee D. Changes associated with quitting

- cigarette smoking: the Framingham Study. *American heart journal* 1975;**90**(3):322-328.
13. John U, Meyer C, Hanke M, Völzke H, Schumann A. Smoking status, obesity and hypertension in a general population sample: a cross-sectional study. *Qjm* 2006;**99**(6):407-415.
 14. Okubo Y, Suwazono Y, Kobayashi E, Nogawa K. An association between smoking habits and blood pressure in normotensive Japanese men: a 5-year follow-up study. *Drug and alcohol dependence* 2004;**73**(2):167-174.
 15. Radi S, Lang T, Lauwers-Cances V, Chatellier G, Fauvel J, Larabi L, De Gaudemaris R. One-year hypertension incidence and its predictors in a working population: the IHPAF study. *Journal of human hypertension* 2004;**18**(7):487-494.
 16. Retnakaran R, Hanley AJ, Connelly PW, Harris SB, Zinman B. Cigarette smoking and cardiovascular risk factors among Aboriginal Canadian youths. *Canadian Medical Association Journal* 2005;**173**(8):885-889.
 17. Mann SJ, James GD, Wang RS, Pickering TG. Elevation of ambulatory systolic blood pressure in hypertensive smokers: a case-control study. *Jama* 1991;**265**(17):2226-2228.
 18. Verdecchia P, Schillaci G, Borgioni C, Ciucci A, Zampi I, Battistelli M, Gattobigio R, Sacchi N, Porcellati C. Cigarette smoking, ambulatory blood pressure and cardiac hypertrophy in essential hypertension. *Journal of hypertension* 1995;**13**(10):1209-1216.
 19. Gambini G, Di Cato L, Pinchi G, Valori C. [24-hour ambulatory monitoring of

- arterial blood pressure and the sympathetic nervous system in hypertensive smokers]. *Giornale italiano di cardiologia* 1997;**27**(11):1153-1157.
20. Berglund G, Wilhelmsen L. Factors related to blood pressure in a general population sample of Swedish men. *Acta medica scandinavica* 1975;**198**(1-6):291-298.
 21. Janzon E, Hedblad B, Berglund G, Engström G. Changes in blood pressure and body weight following smoking cessation in women. *Journal of internal medicine* 2004;**255**(2):266-272.
 22. Lee D-H, Ha M-H, Kim J-R, Jacobs DR. Effects of smoking cessation on changes in blood pressure and incidence of hypertension a 4-year follow-up study. *Hypertension* 2001;**37**(2):194-198.
 23. Dawber TR, Meadors GF, Moore Jr FE. Epidemiological Approaches to Heart Disease: The Framingham Study*. *American Journal of Public Health and the Nations Health* 1951;**41**(3):279-286.
 24. Cryer PE, Haymond MW, Santiago JV, Shah SD. Norepinephrine and epinephrine release and adrenergic mediation of smoking-associated hemodynamic and metabolic events. *New England journal of medicine* 1976;**295**(11):573-577.
 25. Grassi G, Seravalle G, Calhoun DA, Bolla GB, Giannattasio C, Marabini M, Del Bo A, Mancia G. Mechanisms responsible for sympathetic activation by cigarette smoking in humans. *Circulation* 1994;**90**(1):248-253.
 26. Rubin DB. Formal mode of statistical inference for causal effects. *Journal of Statistical Planning and Inference* 1990;**25**(3):279-292.
 27. Hernán M. A definition of causal effect for epidemiological research. *Journal of*

- Epidemiology and Community Health* 2004;**58**(4):265-271.
28. Pearl J. Causal inference in statistics: An overview. *Statistics Surveys* 2009;**3**:96-146.
 29. VanderWeele TJ. Concerning the consistency assumption in causal inference. *Epidemiology* 2009;**20**(6):880-883.
 30. Avin C, Shpitser I, Pearl J. Identifiability of path-specific effects. *Department of Statistics, UCLA* 2005.
 31. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction* Oxford University Press, 2015.
 32. Geneletti S. Identifying direct and indirect effects in a non-counterfactual framework. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2007;**69**(2):199-215.
 33. Didelez V, Dawid P, Geneletti S. Direct and indirect effects of sequential treatments. *arXiv preprint arXiv:1206.6840* 2012.
 34. VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology* 2014;**25**(2):300-306.
 35. Tchetgen EJT, VanderWeele TJ. On identification of natural direct effects when a confounder of the mediator is directly affected by exposure. *Epidemiology (Cambridge, Mass.)* 2014;**25**(2):282.
 36. Taubman SL, Robins JM, Mittleman MA, Hernán MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *International Journal of Epidemiology* 2009;**38**(6):1599-1611.

37. Roger Logan ST, Jessica Young, Sally Picciotto, Miguel A. Hernán. GFORMULA SAS MACRO - Estimates the mean of a dichotomous outcome at end of follow-up under general interventions on time-varying treatments in observational studies using the parametric g-formula.
<http://www.hsph.harvard.edu/causal/software/> 2015.
38. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J. Factors of Risk in the Development of Coronary Heart Disease—Six-Year Follow-up ExperienceThe Framingham Study. *Annals of internal medicine* 1961;**55**(1):33-50.
39. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychological methods* 2010;**15**(4):309.
40. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology* 2009;**20**(1):18-26.
41. Tchetgen Tchetgen EJ. Identification and estimation of survivor average causal effects. *Statistics in medicine* 2014;**33**(21):3601-3628.
42. Tchetgen EJT, Glymour MM, Shpitser I, Weuve J. Rejoinder: to weight or not to weight?: on the relation between inverse-probability weighting and principal stratification for truncation by death. *Epidemiology* 2012;**23**(1):132-137.
43. Robins JM, Wasserman L. Estimation of effects of sequential treatments by reparameterizing directed acyclic graphs. *Proceedings of the thirteenth conference on uncertainty in artificial intelligence: Morgan Kaufmann Publishers Inc.*, 1997;409-420.
44. Young JG, Tchetgen Tchetgen EJ. Simulation from a known Cox MSM using standard parametric models for the g-formula. *Statistics in medicine*

2014;33(6):1001-1014.

45. Van der Laan MJ, Rose S. Targeted learning: causal inference for observational and experimental data Springer Science & Business Media, 2011.

Table 1.1. Summary of covariate models.

Variable	Type of model when used as dependent variable	Functional form when used as predictor
Non-modifiable		
Gender	Not predicted	Indicator
Age	Not predicted	Quadratic linear
Height	Not predicted	Quadratic linear
Education level	Not predicted	Six categories ^a
Occupation	Not predicted	Six categories ^a
Marital status	Not predicted	Three categories ^b
Baseline smoking	Not predicted	Three categories ^b
Modifiable		
SBP	Linear	Quadratic linear
Smoking	Logistic then log-linear ^c	Quadratic linear
Cholesterol	Linear	Quadratic linear
Anti-hypertension drug	Linear	Three categories ^b

^a Education level categories are $\leq 8^{\text{th}}$ grade, some high school, high school graduate, some college, college graduate, and post-graduate. Occupation categories are executive/supervisory, technical, laborer, clerical, sales, and housewife.

^b Marital status categories are single, married, and divorce or widowed. Baseline smoking are smoking, not smoking, and quitting. Anti-hypertension drug are regular use, not use, and sporadic use.

^c zero-continuous variables such as cigarettes per day are predicted in two stages, first a logistic regression on an indicator of whether the variable is nonzero and then a linear regression of the log of the nonzero values.

Table 1.2. Baseline characteristics of eligible participants grouped by former smoking status.

Characteristic	Quitters (n = 183)	Non-smokers (n = 1,174)	Smokers (n = 1,759)	p-value
Male (%)	128 (69.9)	188 (16.0)	1049 (59.6)	<0.001
Age, year (Mean (SD))	50.10 (8.49)	48.90 (8.49)	45.87 (7.96)	<0.001
SBP, mmHg (Mean (SD))	128.53 (17.79)	131.64 (21.54)	125.97 (17.98)	<0.001
BMI (Mean (SD))	25.92 (3.51)	26.22 (4.22)	25.16 (3.68)	<0.001
Chol (Mean (SD))	233.89 (44.08)	232.18 (44.75)	227.54 (42.97)	0.008
Height (Mean (SD))	66.41 (3.46)	62.93 (3.20)	65.46 (3.42)	<0.001
Education (%)				0.019
<High school	65 (35.5)	507 (43.2)	675 (38.4)	

High school	64 (35.0)	420 (35.8)	707 (40.2)	
College or higher	51 (27.9)	227 (19.3)	350 (19.9)	
Missing	3 (1.6)	20 (1.7)	27 (1.5)	
Work (%)				<0.001
Supervisory	58 (31.7)	156 (13.3)	419 (23.8)	
Technical	20 (10.9)	31 (2.6)	130 (7.4)	
Laborer	48 (26.2)	238 (20.3)	517 (29.4)	
Clerical	12 (6.6)	76 (6.5)	109 (6.2)	
Sales	8 (4.4)	19 (1.6)	91 (5.2)	
Housewife	36 (19.7)	647 (55.1)	478 (27.2)	
Missing	1 (0.5)	7 (0.6)	15 (0.9)	
Marital (%)				<0.001
Married	163 (89.1)	935 (79.6)	1580 (89.8)	
Single	10 (5.5)	145 (12.4)	105 (6.0)	
Divorced	10 (5.5)	94 (8.0)	74 (4.2)	

SBP: systolic blood pressure; BMI: body mass index; Chol: cholesterol level;

No one use the anti-hypertensive drugs at the beginning.

**Table 1.3. Estimates of the total effect of smoking 20 cigarettes per day for 10 years
(compared with no smoking) on SBP and BMI.**

Intervention	SBP	Change of SBP	BMI	Change of BMI
	(mm-Hg)	(95% CI)	(kg/m ²)	(95% CI)
No intervention	136.22	0.52 (-0.08, 1.02)	25.83	-0.04 (-0.20, 0.05)
no smoking	135.70	reference	25.87	reference
20 cigarettes/day	136.88	1.18 (-1.01, 3.14)	25.71	-0.16 (-0.37, 0.02)

SBP: systolic blood pressure; BMI: body mass index; CI: confident interval;

Table 1.4. Randomly interventional analogue of total effect, of natural direct effect, and of natural indirect effect for the effect of smoking 20 cigarettes per day for 10 years (compared with no smoking) on SBP, mediated by BMI change over time.

	Estimate	95% CI
$E[Y_{0G0}]$	135.691	134.93, 137.11
$E[Y_{1G0}]$	137.211	135.76, 138.80
$E[Y_{0G1}]$	135.336	134.57, 136.69
$E[Y_{1G1}]$	136.874	135.64, 138.37
rTE	1.18	-0.68, 2.69
rNDE	1.52	-0.25, 2.90
rNIE	-0.34	-0.52, -0.13

rTE: randomly interventional analogue of total effect; rNDE: randomly interventional analogue of natural direct effect; rNIE: randomly interventional analogue of natural indirect effect; CI: confident interval; $E[Y_{0G0}]$, $E[Y_{1G0}]$, $E[Y_{0G1}]$, and $E[Y_{1G1}]$ represent “no smoking with BMI distributed as the BMI under no smoking”, “smoking 20 cigarettes per day with BMI distributed as the BMI under no smoking”, “no smoking with BMI distributed as the BMI under smoking 20 cigarettes per day”, and “smoking 20 cigarettes per day with BMI distributed as the BMI under smoking 20 cigarettes per day”, respectively.

Figure 1.1.Simple model for mediation analysis.

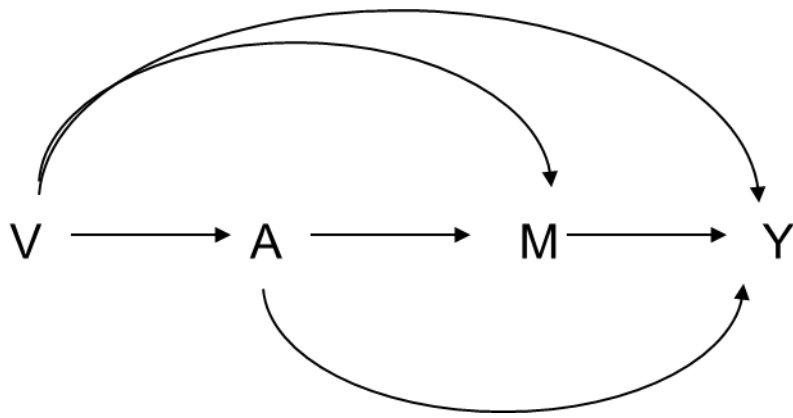


Figure 1.2. Mediation analysis with a mediator-outcome confounder L that is affected by exposure.

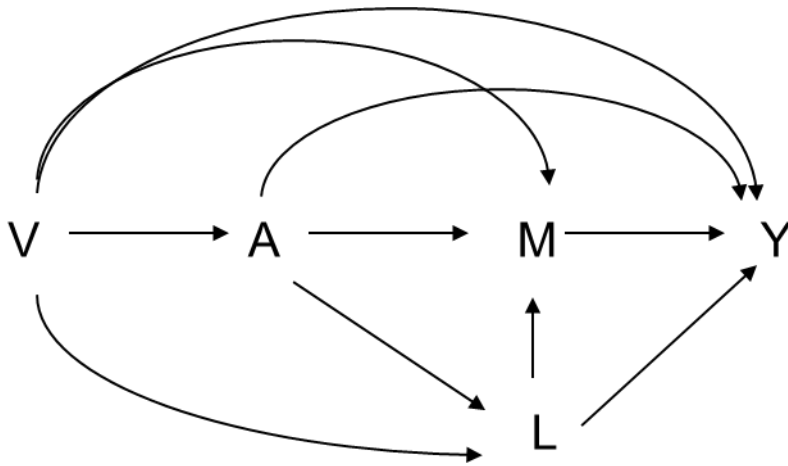
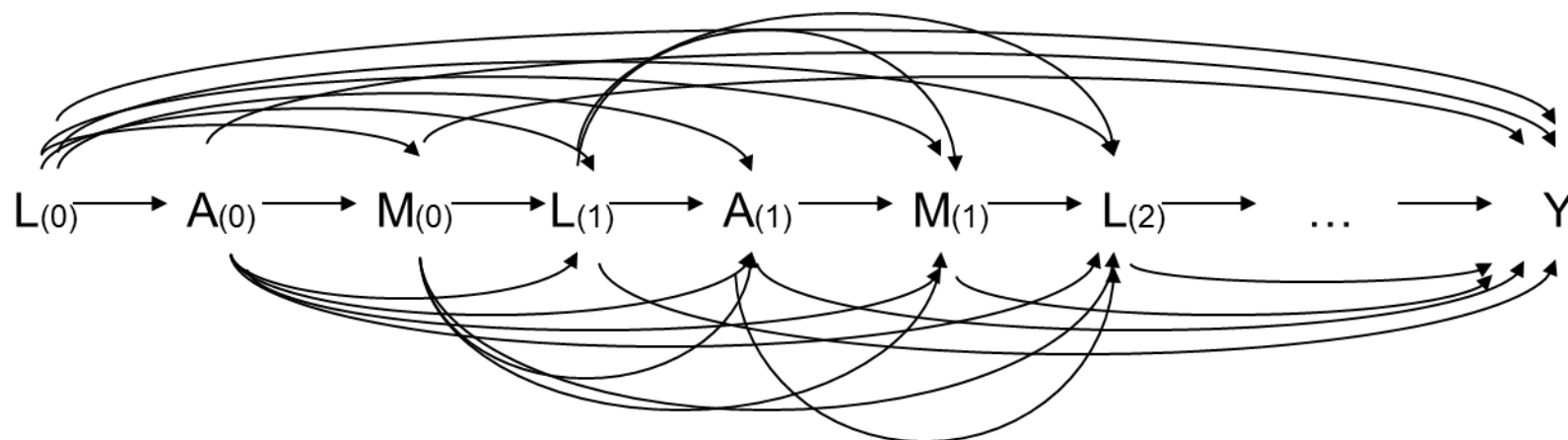


Figure 1.3. Time-varying mediation with ordering of variables of $A(t)$, $M(t)$, $L(t)$, for $t = 0$ to $T-1$.



**Paper 2 Mediation analysis for survival outcome with time-varying exposures,
mediators, and confounders: a case study of the Framingham Heart Study**

Sheng-Hsuan Lin^{1,2}, Eric J. Tchetgen Tchetgen^{1,2}, Jessica Young³, Roger Logan¹, Tyler J.
VanderWeele^{*1,2}

*corresponding author

¹Department of Epidemiology, Harvard School of Public Health, Boston, MA

²Department of Biostatistics, Harvard School of Public Health, Boston, MA

³Department of Population Medicine, Harvard Medical School, Boston, MA

Abstract

This study proposes an approach to conduct mediation analysis for survival data with time-varying exposures, mediators, and confounders. We identify the direct and indirect effects through a survival mediational g-formula and provide the required assumptions. We also provide a feasible parametric approach along with an algorithm and software to estimate these effects. We apply this method to analyze the Framingham Heart Study data to investigate the causal mechanism of smoking on mortality. The risk ratio of smoking 30 cigarettes per day for ten years compared with no smoking on mortality is 2.34 (95 % CI = (1.44, 3.70)). Among the effect, 7.91 % is mediated by coronary artery disease. The survival mediational g-formula constitutes a powerful tool for conducting mediation analysis with longitudinal data.

Introduction

Background

In decomposing the total effect into direct and indirect effects, mediation analysis is essential for investigating pathways or mechanisms in epidemiology and the social sciences. Causal mediation analysis defines both direct and indirect effects based on counterfactual models, extending traditional mediation analysis to settings involving nonlinearities and interaction ^{1,2}. Numerous methodological approaches based on causal mediation analysis have been developed in recent years allowing different outcome scales, including the risk difference, the odds ratio, and the time-to-event data ³⁻⁸. Most of the approaches mentioned above only consider one exposure along with one mediator. When conducting causal mediation analysis with longitudinal data, only one single exposure and mediator ignore the exposures or mediators at other time-points, thus eliminating large quantities of valuable information.

VanderWeele and Tchetgen Tchetgen have proposed the mediational g-formula to address the methodological challenges of causal mediation analysis with time-varying mediators ⁹. This method decomposes the total effect into newly defined direct and indirect effects and therefore moves beyond the limitations of a single exposure and a single mediator. Using this method, time-varying confounders are also adjusted for. Lin et al. proposed a parametric g-computation approach with an available SAS macro to implement this method for the Framingham Heart Study (FHS), investigating the mechanism of smoking on systolic blood pressure mediated by weight change over time ¹⁰. The mediational g-formula allows the outcome of interest as a variable at the end of

follow up. For a survival outcome, Zheng and van der Laan have provided a similar approach with time-varying mediators, one fixed exposure, and baseline confounders ¹¹.

In the current study, we extend the mediational g-formula and Zheng's and van der Laan's technique to offer a generalized setting with survival outcome and time-varying mediators, exposures, and confounders. We define and identify the direct and indirect effects, propose a parametric approach to estimate these effects, and provide an algorithm as well as a corresponding Statistical Analysis System (SAS) macro for practical application. We then apply this method to the FHS data to investigate the effect of smoking on mortality mediated by coronary artery disease (CAD).

The paper is organized as follows: Section 1.2 describes the case study, and section 2 introduces the notations and definitions of the direct and indirect effects of interest. Section 3 presents the non-parametric identification of the direct and indirect effects, and the required assumptions. Section 3 also shows that our formula reduces to the formula provided by Zheng and van der Laan ¹¹ when both an exposure and confounders are not time-varying. Our approach however is applicable with time-varying exposures and confounders. A feasible parametric approach along with an algorithm and software are also provided. Section 4 describes the analytical procedure and provides the estimation results for the FHS data. Section 5 concludes by discussing the strengths and limitations of the study.

Case study for Framingham dataset: smoking, coronary artery disease, and overall mortality

Using the FHS data as an example, our main focus is the effect of smoking on mortality mediated by coronary artery disease (CAD). In the FHS data, smoking and CAD status vary over time. We obtain the estimates of this effect by applying our method to the FHS data and demonstrate the applications of mediation analysis with time-varying covariates (including mediators, exposures, and confounders). We treat smoking status and CAD as exposures and mediators of interest, respectively.

Notations and definitions

Notations

Consider that the exposures, mediators, confounders, and survival outcomes vary over time in longitudinal data with 1 to T measurement. Let $(A(1), \dots, A(T))$, $(M(1), \dots, M(T))$, $(L(1), \dots, L(T))$, and $(S(1), \dots, S(T))$ denote values of time-varying exposures, mediators, confounders, and survival outcomes at periods 1,..., T, with initial baseline confounders, V. The survival outcome at the end of follow up ($S(T)$) is the outcome of interest. $S(t) = 1$ indicates survival at time t; and $S(t) = 0$ indicates death at time t. The causal relationship among these variables is demonstrated in Figure 2.1.

For any variable W and value w, let $W(t_1:t_2) = (W(t_1), W(t_1+1), \dots, W(t_2))$ and $w(t_1:t_2) = (w(t_1), w(t_1+1), \dots, w(t_2))$, while t_1 and t_2 are all positive integers and $t_1 < t_2$. Let $A(1:T) = a(1:T)$ and $A(1:T) = a(1:T)^*$ denote two hypothetical intervention statuses, exposure and non-exposure, respectively. Let $S(t)_{a(1:t)m(1:t)}$ be the counterfactual value of

$S(t)$, given the previous exposures $A(1:t)$ are set to $a(1:t)$ and the previous mediators $M(1:t)$ are set to $m(1:t)$. Let $S(t)_{a(1:t)m(1:t)s(1:t-1)}$ be the counterfactual value of $S(t)$, given the previous exposures $A(1:t)$ are set to $a(1:t)$, the previous mediators $M(1:t)$ are set to $m(1:t)$, and $S(1:t-1)$ are set to $s(1:t-1)$. Here we are conceiving an intervention on survival. Another way forward for this problem is addressed by principal stratification^{12,13}, but this is difficult to the context of mediation. So we assume that the intervention on survival is possible as does prior literature. Let $M(t)_{a(1:t)}$ and $S(t)_{a(1:t)}$ be the counterfactual value of $M(t)$ and $S(t)$, respectively, given $A(1:t)$ are set to $a(1:t)$. Let $G(t)_{a(1:t)}$ denote a random draw from the distribution of the mediator $M(t)_{a(1:t)}$. Let $M(t)_{a(1:t)m(1:t)s(1:t-1)}$ be the counterfactual value of $M(t)$, given the previous exposures, $A(1:t)$, are set to $a(1:t)$, the previous mediators, $M(1:t-1)$, are set to $m(1:t-1)$, and $S(1:t-1)$ is set to $s(1:t-1)$. Since the $S(t)$ is the survival variable, when $S(t) = 0$, the following survival variables (i.e. $S(k)$, while $k \in \{t+1, t+2, \dots, T\}$) are all equal to zero, and the following variables other than $S(k)$ (i.e. $A(k)$, $M(k)$, $L(k)$, while $k \in \{t+1, t+2, \dots, T\}$) are all undefined. In addition, we also make the consistency assumption^{6,14,15}. Under this assumption, $S(t)_{a(1:t)} = S(t)$ and $M(t)_{a(1:t)} = M(t)$ given $A(1:t) = a(1:t)$; $S(t)_{a(1:t)m(1:t)} = S(t)$ given $A(1:t) = a(1:t)$ and $M(1:t) = m(1:t)$; $S(t)_{a(1:t)m(1:t)s(1:t-1)} = S(t)$ given $A(1:t) = a(1:t)$, $M(1:t) = m(1:t)$, and $S(t-1) = s(t-1)$; and $M(t)_{a(1:t)m(1:t-1)s(1:t-1)} = M(t)$, given $A(1:t) = a(1:t)$, $M(1:t-1) = m(1:t-1)$, and $S(t-1) = s(t-1)$.

For convenience of the notation in Section 2.2 and Section 3, we define $M(t)^*$, $G(t)^*$, and $S(t)^*$ sequentially. When time = 1, $M(1)^*$ is defined as $M(1)_{a(1)^*}$, $G(1)^*$ as the random draw of $M(1)^*$, and $S(1)^*$ as $S(1)_{a(1), G(1)^*}$. When time = 2, $M(2)^*$ is defined as $M(2)_{a(1:2)^*, G(1)^*, S(1)^*}$, $G(2)^*$ as the random draw of $M(2)^*$, and $S(2)^*$ as $S(2)_{a(1:2), G(1:2)^*, S(1)^*}$.

We continue this definition process iteratively. For time = t, $M(t)^*$ is defined as $M(t)_{a(1:t)^*, G(1:t-1)^*, S(1:t-1)^*}$, $G(t)^*$ as the random draw of $M(t)^*$, and $S(t)^*$ as $S(t)_{a(1:t), G(1:t)^*, S(1:t-1)^*}$. To conveniently illustrate the distinction of the new definition from the traditional definition, we introduce $M(t)^{**}$ and $S(t)^{**}$ by a similar definition process as $M(t)^*$ and $S(t)^*$. When time = 1, $M(1)^{**}$ is defined as $M(1)_{a(1)^*}$ (which is exactly equal to $M(1)^*$), and $S(1)^{**}$ as $S(1)_{a(1), M(1)^{**}}$ (which replace the $G(1)^*$ by $M(1)^*$ in $S(1)^*$). When time = 2, $M(2)^{**}$ is defined as $M(2)_{a(1:2)^*, M(1)^{**}, S(1)^{**}}$, and $S(2)^{**}$ as $S(2)_{a(1:2), M(1:2)^{**}, S(1)^{**}}$. We again repeat this definition process. For time = t, $M(t)^{**}$ is defined as $M(t)_{a(1:t)^*, M(1:t-1)^{**}, S(1:t-1)^{**}}$, and $S(t)^{**}$ as $S(t)_{a(1:t), M(1:t)^{**}, S(1:t-1)^{**}}$. Compared with $S(t)^*$, we define $S(t)^{**}$ by using $M(t)^{**}$ directly, rather than its random draw $G(t)^{**}$.

Since $M(t)$ is undefined when $S(t-1) = 0$, under the counterfactual model, $M(t)_{s(t-1)=0}$ is also undefined. Consequently, $S(t)_{m(t), s(t-1)=1}$ is undefined when $m(t)$ is undefined. Since $S(t)$ is always equal to zero when $S(t-1) = 0$, $S(t)_{m(t), s(t-1)=0}$ is always equal to zero even when $m(t)$ is undefined. $M(t)^*$ is undefined if and only if $S(t-1)^* = 0$. Therefore, $S(t)^*$ is still defined even when $M(t)^*$ is undefined since in that condition, $S(t-1)^*$ should be zero, leading to $S(t)^* = 0$. Similarly, $S(t)^{**}$ is always defined, too.

Definitions of mediation parameter and the direct and indirect effects

Aiming to investigate the extent of the effect of exposures over time ($A(1:T)$) on the survival outcome at the end of follow up ($S(T)$), through mediators over time ($M(1:T)$), mediation analysis decomposes the overall effect into a direct effect (the effect not through the mediators) and an indirect effect (the effect through the mediators). The

direct and indirect effects are traditionally represented by the natural direct effect (NDE) and the natural indirect effect (NIE), respectively; NDE and NIE are mathematically defined as $E[S(T)_{a(1:T),M(1:T)_{a^*}} - S(T)_{a(1:T)^*,M(1:T)_{a(1:T)^*}}|v]$ and $E[S(T)_{a(1:T),M(1:T)_{a(1:T)}} - S(T)_{a(1:T),M(1:T)_{a(1:T)^*}}]$, respectively^{1,2,6}. For simplicity of expression, we define the mediation parameter $\Phi(a(1:T), a(1:T)^*)$ as $E[S(T)_{a(1:T),M(1:T)_{a(1:T)^*}}]$. Thus, we can rewrite the NDE and NIE as $\Phi(a(1:T), a(1:T)^*) - \Phi(a(1:T)^*, a(1:T)^*)$ and $\Phi(a(1:T), a(1:T)) - \Phi(a(1:T), a(1:T)^*)$, respectively. This definition of NIE captures all effects of $A(1:T)$ on $S(T)$ through $M(1:T)$.

The direct and indirect effects can also be represented by interventional direct effect (IDE) and interventional indirect effect (IIE), of which the mediation parameter is defined as $E[S(T)_{a(1:T),G(1:T)_{a(1:T)^*}}]$. In some literatures, IDE and IIE were also named randomly interventional analogues of natural direct effect and of natural indirect effect, respectively^{9,10,16}. When the time-varying confounders exist, both IDE and IIE can be identified by the empirical dataset, but NDE and NIE cannot.

The definition of NDE and NIE presents a major problem for individuals when the exposures benefit the survival status at time $T-1$, not through the mediators (mathematically expressed as $S(T-1)_{a(1:T),M(1:T-1)_{a(1:T)^*}} = 1$ and $S(T-1)_{a(1:T)^*,M(1:T-1)_{a(1:T)^*}} = 0$). Based on this mathematical expression, $M(1:T)_{a(1:T)^*}$ (defined as $M(1:T)_{a(1:T)^*,M(1:T-1)_{a(1:T)^*},S(T-1)_{a(1:T)^*,M(1:T-1)_{a(1:T)^*}}$) is equal to $M(1:T)_{a(1:T)^*,M(1:T-1)_{a(1:T)^*},S(T-1)=0}$, which is undefined. Consequently, the mediation parameter $\Phi(a(1:T), a(1:T)^*)$ ($= S(T)_{a(1:T),M(1:T)_{a(1:T)^*}}$)

$S(T)_{a(1:T), M(1:T)_{a(1:T)^*, S(1:T-1)_{a(1:T), M(1:T)_{a(1:T)^*}}$ by definition) is also undefined. In addition, NDE and NIE, which are both defined by $\Phi(a(1:T), a(1:T)^*)$, are also undefined. Both IDE and IIE are also undefined since G is a random draw of M .

One way to address the undefined NDE and NIE (as well as IDE and IIE) is changing the definition of the mediation parameter (Zheng and van der Laan, 2012). In Section 2.1, we have shown that both $S(t)^*$ is always defined. The alternative mediation parameter, $\Psi(a(1:T), a(1:T)^*)$, is defined as $E[S(T)^*]$. IDE and IIE are then defined as $\Psi(a(1:T), a(1:T)^*) - \Psi(a(1:T)^*, a(1:T)^*)$ and $\Psi(a(1:T), a(1:T)) - \Psi(a(1:T), a(1:T)^*)$, respectively. Since $S(t)^{**}$ is also always defined, we can use the $E[S(T)^{**}]$ to define mediation parameter, as well as NDE and NIE. Since time-varying confounders cannot be ignored in settings with time-varying exposures and mediators, we consider $E[S(T)^*]$ rather than $E[S(T)^{**}]$ to define mediation parameter $\Psi(a(1:T), a(1:T)^*)$ and try to identify and estimate IDE and IIE in the current study.

These definitions (by $S(T)^*$ or $S(T)^{**}$) are based on the path-specific effect^{9,11,17}. Under the traditional definition, an indirect effect consists of two groups of path-specific effects: first, the path of exposures affecting the mediators through earlier survival history, and second, the path of exposures affecting the mediators not through earlier survival history. Under our definition for the alternative mediation parameter ($\Psi(a(1:T), a(1:T)^*)$), the indirect effect includes the second path-specific effect, and the first path-specific effect is included in the direct effect.

We now illustrate the distinction between traditional definition and our definition by taking the simplest setting as an example: given $T = 2$, and V , $A(2)$, $L(1)$, and $L(2)$ are all

empty (Figure 2.2). Since the mediation parameter defined by $E[S(T)^*]$ or $E[S(T)^{**}]$ captures the similar path. For heuristic concerning the effects, we consider $E[S(t)^{**}]$ as an analogue of $E[S(T)^*]$. According to traditional definition, NIE includes (a) $A \rightarrow M(2) \rightarrow S(2)$, (b) $A \rightarrow M(1) \rightarrow M(2) \rightarrow S(2)$, (c) $A \rightarrow M(1) \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$, (d) $A \rightarrow M(1) \rightarrow S(1) \rightarrow S(2)$, and (e) $A \rightarrow M(1) \rightarrow S(2)$; NDE includes (a) $A \rightarrow S(1) \rightarrow S(2)$, (b) $A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$, and (c) $A \rightarrow S(2)$. From the eight possible path-specific effects, $A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$ is traditionally included as part of NIE since $M(2)$ is involved. Under our definition, $A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$ is included in NDE since the exposure affects survival history ($S(1)$) first, not through the earlier mediator ($M(1)$). The path ($A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$) can be measured by $E[S(2)_{a,M(1)_{a*},S(1)_{a,M(1)_{a*}},M(2)_{a*,M(1)_{a*},S(1)_{a,M(1)_{a*}}}] - E[S(2)_{a,M(1)_{a*},S(1)_{a,M(1)_{a*}},M(2)_{a*,M(1)_{a*},S(1)_{a*,M(1)_{a*}}}]$ (i.e., $E[S(2)^{**}] - \Phi(a,a^*)$; here, we use $E[S(2)^{**}]$ rather than $E[S(2)^*]$ for heuristic simplicity). When exposure affects survival status at time 1 beneficially (mathematically expressed as $(S(1)_{a,M(1)_{a*}}, S(1)_{a*}) = (1,0)$), the $\Phi(a,a^*) (= E[S(2)_{a,M(1)_{a*},S(1)=1,M(2)_{a*,M(1)_{a*},S(1)=0}])$ is undefined, and consequently, the path, $A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$, is undefined. On the other hand, when exposure has monotonically harmful effect on survival status at time 1 (mathematically expressed as $(S(1)_{a,M(1)_{a*}}, S(1)_{a*}) \in \{(1,1), (0,0), (0,1)\}$), the path is always equal to zero (when $S(1)_{a,M(1)_{a*}} = 0$, $E[S(2)^{**}] = E[S(2)_{a,M(1)_{a*},S(1)=0,M(2)_{a*,M(1)_{a*},S(1)=0}] = 0$ and $\Phi(a,a^*) = E[S(2)_{a,M(1)_{a*},S(1)=0,M(2)_{a*,M(1)_{a*},S(1)_{a*,M(1)_{a*}}}] = 0$; when $S(1)_{a,M(1)_{a*}} = 1$ and $S(1)_{a*} = 1$, $E[S(2)^{**}]$ and $\Phi(a,a^*)$ are all equal to $E[S(2)_{a,M(1)_{a*},S(1)=1,M(2)_{a*,M(1)_{a*},S(1)=1}]$). In conclusion, the difference between the traditional and alternative mediation parameters is

always undefined or zero. Thus, the path-specific effect is a reasonable alternative definition for the direct and indirect effects.

Identification and estimation of survival mediational g-formula

Assumptions and identification for survival mediational g-formula

For identifying the mediation parameter, $\Psi(a(1:T), a(1:T)^*)$, we should make the following four, sequential no unmeasured confounding assumptions for $t = 1, 2, \dots, T$:

1. $S(T)_{a(1:t), m(1:t), s(1:t-1)=1, G(t+1:T)^*, S(t:T-1)^*} \perp A(t)|v, A(1:t-1)=a(1:t-1), M(1:t-1)=m(1:t-1), L(1:t-1), S(1:t-1)^*=1, G(1:t)^* = m(1:t)$ (no unmeasured exposure-outcome confounding conditional on the past covariates, while \perp indicates independence)
2. $S(T)_{a(1:t), m(1:t), s(1:t-1)=1, G(t+1:T)^*, S(t:T-1)^*} \perp M(t)|v, A(1:t)=a(1:t), M(1:t-1)=m(1:t-1), L(1:t-1), S(1:t-1)^*=1, G(1:t)^*=m(1:t)$ (no unmeasured mediator-outcome confounding conditional on the past covariates)
3. $M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} \perp S(t-1)|v, A(1:t-1), M(1:t-1), L(1:t-1), S(1:t-2)=1$ (no unmeasured mediator-previous survival confounding conditional on the past covariates)
4. $M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} \perp A(t)|A(1:t-1), M(1:t-1), L(1:t-1), S(1:t-1)=1, v$ (no unmeasured exposure-mediator confounding conditional on the past covariates)

It is worth noting that an unmeasured confounder of survival statuses at different time points is allowed. None of the assumptions (even assumption 3) will be violated. We

will illustrate this situation in Appendix 2.4. Under the four assumptions, the mediation parameter $\Psi(a(1:T), a(1:T)^*)$ can be identified as $Q(a(1:T), a(1:T)^*)$

=

$$\begin{aligned} & \sum_{v, m(1:T)} \sum_{l(1:T)} \prod_{t=1}^T E[S(t) | a(1:t), m(1:t), l(1:t), S(t-1) = 1, v] \times \\ & \prod_{t=1}^{T-1} P(l(t) | a(1:t), m(1:t), l(1:t-1), S(t-1) = 1, v) \times \\ & \sum_{l'(1:T-1)} \prod_{t=1}^T \Pr(m(t) | a^*(1:t), m(1:t-1), l'(1:t-1), S(1:t-1) = 1, v) \times \\ & \Pr(l'(t-1) | a^*(1:t-1), m(1:t-1), l'(1:t-2), S(1:t-2) = 1, v) \Pr(v) \end{aligned}$$

The $Q(a(1:T), a(1:T)^*)$ can also be expressed by the counting process notation, $N(t)$, and the continuous time-varying mediators and confounders as follows.

$$Q(a(1:T), a(1:T)^*)$$

=

$$\begin{aligned} & \int_{m(1:T), l(1:T)} \prod_{t=1}^T E[1 - N(t) | a(1:t), m(1:t), l(1:t), N(t-1) = 0, v] \times \\ & \prod_{t=1}^{T-1} f_{L(t)}(l(t) | a(1:t), m(1:t), l(1:t-1), N(t-1) = 0, v) dl(1:T) \times \\ & \int_{l'(1:T-1)} \prod_{t=1}^T f_{M(t)}(m(t) | a^*(1:t), m(1:t-1), l'(1:t-1), N(1:t-1) = 0, v) \times \\ & f_{L'(t-1)}(l'(t-1) | a^*(1:t-1), m(1:t-1), l'(1:t-2), N(1:t-2) = 0, v) dm(1:T) dl'(1:T-1) \end{aligned}$$

The proof is provided in Appendix 2.1 and 2.3. We refer to this final expression $Q(a(1:T), a(1:T)^*)$ as the survival mediational g-formula (sMGF). Consequently, the IDE and IIE can be identified non-parametrically by the following equations:

$$IDE = Q(a(1:T), a(1:T)^*) - Q(a(1:T)^*, a(1:T)^*)$$

$$IIE = Q(a(1:T), a(1:T)) - Q(a(1:T), a(1:T)^*)$$

Intuitively, the first part

$(\sum_{m(1:T)} \sum_{l(1:T)} \prod_{t=1}^T E[S(t)|a(1:t), m(1:t), l(1:t), S(t-1) = 1, v] \times \prod_{t=1}^{T-1} P(l(t)|a(1:t), m(1:t), l(1:t-1), S(t-1) = 1, v))$ is the g-formula when outcome is a survival variable without censoring (Robins, 1982; Robins et. al., 2004; Taubulin et. al. 2009). The second part

$(\sum_{l'(1:T-1)} \prod_{t=1}^T \Pr(m(t)|a^*(1:t), m(1:t-1), l'(1:t-1), S(1:t-1) = 1, v) \times \Pr(l'(t-1)|a^*(1:t-1), m(1:t-1), l'(1:t-2), S(1:t-2) = 1, v))$ is the joint distribution of mediators, $M(1:T)$, given all exposures are set to $a(1:T)^*$ for survivors.

When mediators are empty, the survival mediational g-formula reduces to the following form:

$$\begin{aligned} E[S(T)_{a(1:T)}] = & \sum_{l(1:T)} \prod_{t=1}^T E[S(t)|a(1:t), l(1:t), S(t-1) = 1, v] \\ & \times \prod_{t=1}^{T-1} P(l(t)|a(1:t), l(1:t-1), S(t-1) = 1, v) \end{aligned}$$

which is the standard g-formula. Thus we can conclude that the survival mediational g-formula is a generalized form of g-formula, similar to the mediational g-formula ⁹.

When $S(T-1)$ is always equal to one, this expression reduces to

$$\begin{aligned} & \sum_{m(1:T)} \sum_{l(1:T)} E[S(T)|a(1:T), m(1:T), l(1:T), v] \\ & \times \prod_{t=1}^{T-1} P(l(t)|a(1:t), m(1:t), l(1:t-1), v) \\ & \times \sum_{l'(1:T-1)} \prod_{t=1}^T \Pr(m(t)|a^*(1:t), m(1:t-1), l'(1:t-1), v) \\ & \times \Pr(l'(t-1)|a^*(1:t-1), m(1:t-1), l'(1:t-2), v) \end{aligned}$$

which is the exact mediational g-formula provided by VanderWeele and Tchetgen Tchetgen's work ⁹, given that the outcome of interest (denoted as Y) is survival status at the end of follow up ($S(T)$).

If neither the confounders and exposure are not time-varying (i.e. $A(2:T)$ and $L(1:T)$ are empty) (Figure 2.3), the survival mediational g-formula reduces to

$$\begin{aligned} & \sum_{m(1:T)} \prod_{t=1}^T E[S(t)|a, m(1:t), S(t-1) = 1, v] \\ & \times \Pr(m(t)|a^*, m(1:t-1), S(t-1) = 1, v) \end{aligned}$$

which is exactly equal to Zheng's and van der Laan's expression (equation 6 and equation 24) ¹¹. However, our survival mediational g-formula can still be used, unlike the approach of Zheng and van der Laan, even if the exposures and confounders are time-varying.

Furthermore, under the monotonicity assumption:

$$S(t)_{a,m(1:t-1)} \leq S(t)_{a^*,m(1:t-1)} \text{ for all individuals where } t = 1, 2, \dots, T-1,$$

the traditional definition of mediation parameter $\Phi(a(1:T), a(1:T)^*)$ can be defined and identified by the following expression:

$$\begin{aligned} & \sum_{m(1:T)} \prod_{t=1}^T E[S(t)|a(1:t), m(1:t), S(t-1) = 1, v] \\ & \times \Pr(m(t)|a(1:t)^*, m(1:t-1), S(t-1) = 1, v) \end{aligned}$$

The proof is provided in Appendix 2.2. This expression is also the special case of the sMGF when both $A(2:T)$ and $L(1:T)$ are empty. In other words, given time-varying confounders do not exist and exposures have monotonically harmful effect on survival, the alternative definitions of IDE and IIE are the same as the traditional definitions of NDE and NIE, respectively (i.e., the difference between the traditional and alternative mediation parameters is equal to zero).

Parametric approach for survival mediational g-formula

VanderWeele and Tchetgen Tchetgen described how to use marginal structure models (MSM) to estimate the mediational g-formula function⁹, which can be inefficient and does not work well with continuous exposures and mediators. Lin et al have proposed a parametric approach for mediational g-formula¹⁸. Because of strong modeling assumptions for outcome and all time-varying covariate modeling, a parametric approach can be more efficient than the MSM approach. We extend this approach to estimate survival mediational g-formula, developing a corresponding algorithm and a SAS macro. Since the survival mediational g-formula is a generalized form of the g-formula, we create the algorithm and the macro for the survival mediational g-formula based on the framework of the g-formula macro (<http://www.hsph.harvard.edu/causal>). First, we specify parametric regression models for the distribution of the time-varying exposures, mediators, confounders, and survival variables. For each model, we include former covariates as the independent variables to try to eliminate confounding and to ensure that the four assumptions (mentioned in Section 3.1) hold. By fitting these models with data, maximal likelihood estimates for all parameters can be obtained. Finally, these estimates are substituted for the parameters in the survival mediational g-formula, deriving consistent estimates for IDE and IIE based on Monte-Carlo simulation.

The algorithm is as follows:

- (1) Fit parametric models for the observed data: for times $t \geq 1$, fit parametric models for the joint density of the confounders, exposures, mediators, and survival status at t given the measured past.

(2) Estimate the joint distribution of time-varying mediators under time-varying exposure interventions $a(1:T)$ and $a(1:T)^*$ given everyone survives.

(2a) Set baseline ($t = 1$) covariates to the observed values for subject i . Recursively, for each time $t = 1, \dots, T$ and each subject $i = 1, \dots, n$:

(2a.i) For $t \geq 1$, generate time t confounders, exposure, and mediator based on the estimated model coefficients of (1a) and previously generated covariates under the time-varying exposure intervention $a(1:t-1)$.

(2a.ii) Assign time t exposure under the intervention $a(1:T)$.

(2b) For each $t = 1, \dots, T$, randomly permute the n values of the joint mediators assigned under intervention $a(1:T)$ in (2a). For each t , save this permutation for use in (3) below (we obtain the random draw of the survivor mediator under exposure, $G(1:T)_{a(1:T), s(T-1)=1}$, in this step).

(2c) Repeat (2a) replacing intervention $a(1:T)$ with $a(1:T)^*$.

(2d) Repeat (2b) replacing intervention $a(1:T)$ with $a(1:T)^*$ (we obtain the random draw of the survivor mediator under non-exposure, $G(1:T)_{a(1:T)^*, s(T-1)=1}$, in this step).

(3) Estimate $Q(a(1:T), a(1:T))$, $Q(a(1:T)^*, a(1:T))$, $Q(a(1:T), a(1:T)^*)$ and $Q(a(1:T)^*, a(1:T)^*)$ by repeating the following for each $(a(1:T)_1, a(1:T)_2) = (a(1:T), a(1:T))$, $(a(1:T), a(1:T)^*)$, $(a(1:T)^*, a(1:T))$, and $(a(1:T)^*, a(1:T)^*)$:

(3a) Recursively for each time $t = 1, \dots, T$ and each subject $i = 1, \dots, n$:

(3a.i) Repeat (2a.i) but replacing “time-varying exposure intervention $a(1:T)$ through $T-1$ ” with the joint “time-varying exposure and mediator intervention $(a(1:T)_1, G(1:T)_{a(1:T)_2, S(T-1)=1})$ ”.

(3a.ii) Assign the time t mediator as the i^{th} component of the permuted vector for time t from (2b) (if $a(1:T)_2 = a(1:T)$) or (2d) (if $a(1:T)_2 = a(1:T)^*$).

(3a.iii) Assign time t exposure under the intervention $a(1:T)_1$.

(3b) Simulate the outcome given each of the $i = 1, \dots, n$ histories based on the estimated model coefficients of (1b) and the histories generated in (3a).

(3c) Estimate $Q(a(1:T)_1, a(1:T)_2)$ as the mean over the n simulated outcomes in (3b).

Estimates of the IDE and IIE are then calculated from the estimates of the four $Q(a(1:T)_1, a(1:T)_2)$ in (3). 95% confidence intervals are calculated based on repeating the above algorithm in 500 bootstrap samples of the original n observations. This algorithm can be implemented with the `mgformula` macro, freely accessible with documentation at <http://www.hsph.harvard.edu/causal/software/>.

Analysis of Framingham Heart Study data

In this section, we apply the survival mediational g -formula to the Framingham Heart Study (FHS) dataset to investigate the causal mechanism of smoking on overall

mortality mediated by CAD. FHS is a longitudinal cohort study beginning in Framingham, Massachusetts in 1948. The original cohort consisted of 5,209 participants between 30 and 62 years old without any symptom of cardiovascular disease (CVD) at baseline. All the participants underwent examinations at the beginning of the study and every two years during the follow up. For each examination, potential CVD risk factors including socio-demographic data, lifestyle characteristics, detailed medical history, physical examination data(1:T), a(1:T)nd blood samples were collected. Further details on the design of FHS are available elsewhere^{19,20}.

Exam 3 is specified as the first exam and exams 1 and 2 as pre-baseline covariates to allow lag predictive models. To mitigate selection bias, we focus on only ten year follow up (i.e. exam 3 to exam 7; the total exam number $T = 5$) to reduce the proportion of death or loss to follow up. Four exclusion criteria are listed below: (1) loss to follow up during the period before exam 7 (the end of follow up); (2) no record at baseline on weight, height, smoking status, former smoking history, systolic blood pressure (SBP), or total cholesterol level; (3) diagnosis of diabetes, cancer, or CVD at baseline; and (4) value for smoking status or BMI missing more than once. After these exclusions, 3,116 participants are eligible for analysis. For simplicity, we now refer to the original FHS exams 3,..., 7 as exams 1,..., 5.

The smoking status at all five exams, measured as self-reported average number of cigarettes smoked per day, are exposures of interest ($A(1:5)$); mortality at the end of follow up is the outcome of interest ($S(5)$); and the CAD status at all exams are the mediators of interest ($M(1:5)$). For missing smoking status for once, we carry forward the last observed smoking status for one exam period only. We considered "smoking 30

cigarettes per day" and "no smoking" at all exams as two hypothetical intervention statuses, $A(1:T) = a(1:T)$ and $A(1:T) = a^*(1:T)$. Time-varying covariates $L(1:T)$ include the exam number, the systolic blood pressure (mm/hg), body mass index (kg/mm), and the usage of antihypertensive drugs. Baseline covariates V include gender and age (years).

The parametric g-formula is used to estimate the total effect of smoking 30 cigarettes per day (v.s. no smoking) on mortality at exam 5, by the g-formula SAS macro. The survival mediational g-formula is applied to conduct mediation analysis with time-varying exposures, mediators, and confounders by the mGFORMULA SAS macro. For conducting the survival mediational g-formula, we specify model for the distribution of time-varying exposures, mediators, confounders, and survival variables at each time point. We use current covariates and covariates at one period back (one lagged model) as the predictors. Specifically, for $t = 1, 2, \dots, 5$, we regress $S(t)$ on $A(t)$, $M(t)$, $L(t)$, $A(t-1)$, $M(t-1)$, and $L(t-1)$; regress $L(t)$ on $A(t)$, $M(t)$, $A(t-1)$, $M(t-1)$, and $L(t-1)$; regress $M(t)$ on $A(t)$, $A(t-1)$, $M(t-1)$, and $L(t-1)$; and regress $A(t)$ on $A(t-1)$, $M(t-1)$, and $L(t-1)$. All analyses are conducted using SAS 9.4 (Cary, NC).

The mortalities among the original group, a hypothetical group if everyone did not smoke for 10 years, and a hypothetical group if everyone smoked 30 cigarettes per day for 10 years, are 4.52 %, 3.37 %, and 7.87 %, respectively (Table 2.1). The risk ratio (RR) is 2.34 (95 % CI = (1.44, 3.70)), which is close to the result calculated by mGFORMULA macro (RR = 2.30; 95 % CI = (1.36, 2.88)). In additive scale, smoking increases mortality by 3.96% directly not through changing CAD, and by 0.34 % through changing CAD. The proportion of the total effect explained by the mediation of CAD is 7.91 % (95 % CI = (1.36, 19.32)) (Table 2.2).

Discussion

This is the first paper providing a generalized framework for causal mediation analysis with a survival outcome and with time-varying exposures, mediators, and confounders. We provide identification for the direct and indirect effects and the required assumptions, using a parametric approach for the point and interval estimation. Based on the g-formula macro, we have developed a feasible algorithm and a mGFORMULA SAS macro, which can also be used for a continuous or binary outcome ^{9,10}. Similar to the g-formula macro, we use the Monte-Carlo simulation and bootstrapping for point and interval estimation, respectively. Since the estimates approximate maximal likelihood estimates, the efficiency of this estimation may be better than the IPW estimation provided the regression models are all correctly specified. Like other simulation-based methods ²¹, our approach has the advantage of allowing for very flexible models, such as quadratic model.

The survival mediational g-formula is currently the only method for investigating the mechanisms of an effect on a survival outcome with time-varying exposures, mediators, and confounders. Under traditional techniques for causal mediation analysis with survival outcome, only one mediator is allowed, and this mediator should occur immediately after the exposure to ensure that the identification assumption holds. Thus traditional techniques do not adequately capture the indirect effect given that the mediators vary over time ^{22,23}. When both the exposure and confounders are fixed, our formula also reduces to Zheng's and van der Laan's formula ¹¹, but our formula can be

used more generally, in addition to the mediators, even when the exposures and confounders vary over time.

Several limitations of this study should be noted. First, our path-specific definition is different from the traditional definition for direct and indirect effect. This is inevitable for causal mediation analysis with time-varying mediators because the NDE/NIE (as well as IDE/IIE) cannot be defined. Because the difference of traditional and alternative definitions is undefined or zero, it is reasonable to use the alternative definition. Second, the outcome only focuses on survival probability at the end of follow up. Extending the outcome model to different survival models such as the Cox proportional hazard model or the accelerated failure time model should be developed for future studies. Third, the non-ignorable drop-out has not been adjusted for here. We restricted our follow up to ten years to partially address this selection bias. The relatively low proportion of loss to follow up ($< 20\%$) perhaps partially mitigates this problem. Our method is also sensitive to the violation of model misspecification, which is the trade-off in obtaining efficient estimates with a parametric approach. However, in our approach, the allowance for using very flexible models (including splines) mitigates this issue. Finally, the analysis is subject to potential violation of the confounding assumptions. Future research could develop sensitivity analysis techniques for violation of these assumptions.

Conclusion

The survival mediational g-formula serves as a powerful and useful tool for mediation analysis with longitudinal data. When the outcome of interest is survival

variable, researchers can apply our method to disentangle the complicated causal mechanisms arising from time-varying mediators, exposures, and confounders.

Reference

1. Pearl J. Direct and indirect effects. Proceedings of the seventeenth conference on uncertainty in artificial intelligence: Morgan Kaufmann Publishers Inc., 2001;411-420.
2. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 1992;143-155.
3. Tchetgen EJT, Shpitser I. Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis. *The Annals of Statistics* 2012;**40**(3):1816-1845.
4. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological methods* 2013;**18**(2):137.
5. van der Laan MJ, Petersen ML. Direct effect models. *The international journal of biostatistics* 2008;**4**(1).
6. VanderWeele T, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface* 2009;**2**:457-468.
7. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology (Cambridge, Mass.)* 2010;**21**(4):540.
8. VanderWeele TJ, Vansteelandt S. Odds ratios for mediation analysis for a

- dichotomous outcome. *American journal of epidemiology* 2010;**172**(12):1339-1348.
9. VanderWeele TJ, Tchetgen Tchetgen E. Mediation Analysis with Time-Varying Exposures and Mediators. 2014.
 10. Lin S-H, Young J, Tchetgen Tchetgen E, VanderWeele TJ. Parametric mediational g-formula approach to mediation analysis with time-varying exposures, mediators and confounders: an application to smoking, weight, and blood pressure. *Epidemiology (accepted)* 2015.
 11. Zheng W, van der Laan MJ. Causal mediation in a survival setting with time-dependent mediators. 2012.
 12. Tchetgen Tchetgen EJ. Identification and estimation of survivor average causal effects. *Statistics in medicine* 2014;**33**(21):3601-3628.
 13. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics* 2002;**58**(1):21-29.
 14. Pearl J. Causal inference in statistics: An overview. *Statistics Surveys* 2009;**3**:96-146.
 15. VanderWeele TJ. Concerning the consistency assumption in causal inference. *Epidemiology* 2009;**20**(6):880-883.
 16. VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology* 2014;**25**(2):300-306.
 17. Avin C, Shpitser I, Pearl J. Identifiability of path-specific effects. *Department of Statistics, UCLA* 2005.

18. Lin S-H, Tchetgen Tchetgen E, VanderWeele TJ. Mediation analysis with survival outcome and time-varying exposures, mediators, and confounders: a case study of the Framingham Heart Study. *submitting* 2015.
19. Dawber TR, Meadors GF, Moore Jr FE. Epidemiological Approaches to Heart Disease: The Framingham Study*. *American Journal of Public Health and the Nations Health* 1951;**41**(3):279-286.
20. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J. Factors of Risk in the Development of Coronary Heart Disease—Six-Year Follow-up ExperienceThe Framingham Study. *Annals of internal medicine* 1961;**55**(1):33-50.
21. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychological methods* 2010;**15**(4):309.
22. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction* Oxford University Press, 2015.
23. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology* 2009;**20**(1):18-26.

Figure 2.1. Time-varying mediation with variables of $A(t)$, $M(t)$, $L(t)$, and $S(t)$, for $t = 1$ to T .

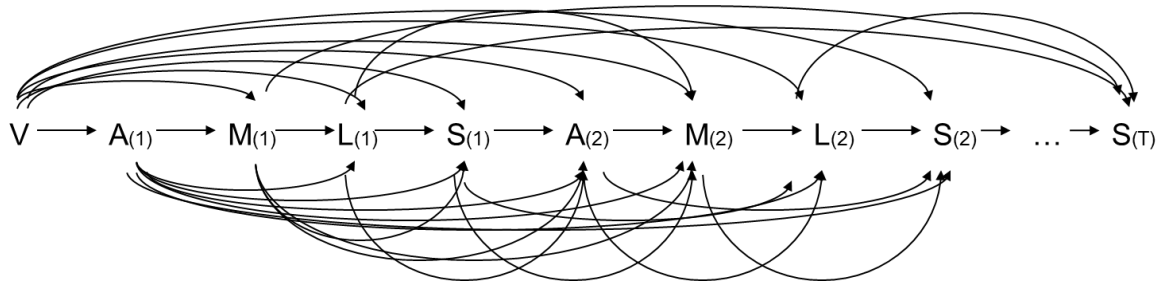


Figure 2.2. Simple model for survival mediation analysis with fixed exposure and time-varying mediators with 2 timepoints.

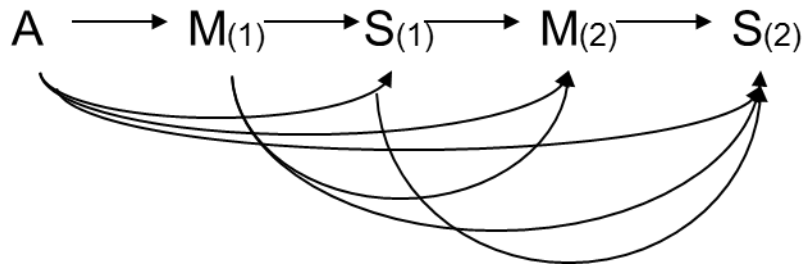


Figure 2.3. Simple model for survival mediation analysis with fixed exposure and time-varying mediators.

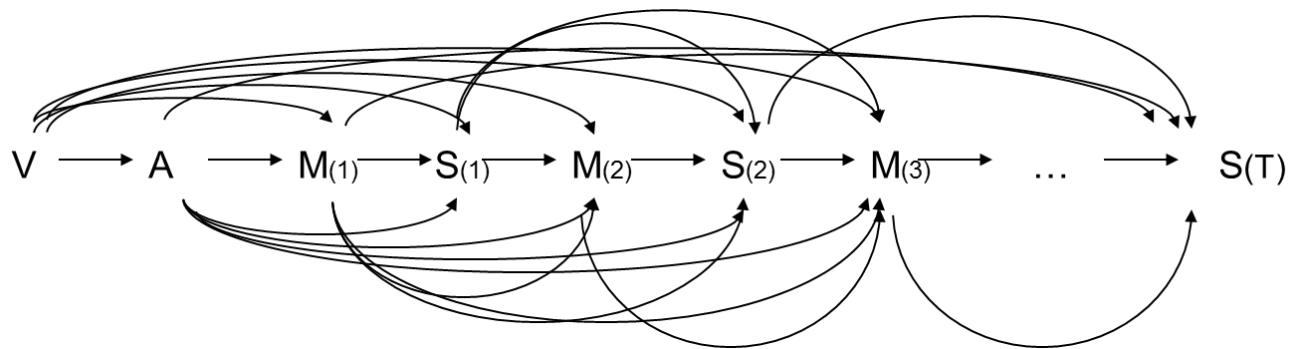


Table 2.1. Estimates of the overall effect of smoking 30 cigarettes per day for 10 year (compared with no smoking) on mortality.

Intervention	Mortality	Difference	95% CI	Ratio	95% CI
	(%)				
No intervention	4.52	1.15	0.58, 2.09	1.34	1.16, 1.69
No smoking	3.37	Ref.		Ref.	
30 cigarettes/day	7.87	4.90	1.59, 8.18	2.34	1.44, 3.70

CI: confident interval.

Table 2.2. Mediation analysis for the effect of smoking 30 cigarettes per day for 10 years (compared with no smoking) on overall mortality, mediated by cardiovascular disease.

Coronary disease		
	Estimate	95% CI
Additive scale		
Total effect	4.3	1.37, 6.30
Direct effect	3.96	1.22, 6.06
Indirect effect	0.34	0.05, 0.96
Multiplicative scale		
Total effect	2.30	1.36, 2.88
Direct effect	2.20	1.33, 2.70
Indirect effect	1.05	1.01, 1.12
Proportion mediated (%)	7.91	1.36, 19.32
CI: confident interval.		

Paper 3 Interventional Approach for Path-Specific Effects

Sheng-Hsuan Lin^{1,2,*}, Tyler J. VanderWeele^{1,2}

*: corresponding author

¹Department of Epidemiology, Harvard School of Public Health, Boston, MA

²Department of Biostatistics, Harvard School of Public Health, Boston, MA

Abstract

Standard causal mediation analysis decomposes the total effect into a direct effect and an indirect effect in settings with only one single mediator. Under the settings with multiple mediators, all mediators are often treated as one single block of mediators. The effect mediated by a certain combination of mediators, i.e. path-specific effect (PSE), is not always identifiable without making strong assumptions. In this paper, the authors propose a method, defining a randomly interventional analogue of PSE (rPSE), which can always be non-parametrically identified under assumptions of no unmeasured confounding. This method also allows settings with mediators dependent on each other, interaction, and mediator-outcome confounders which are affected by exposure. In addition, under linearity and no-interaction, our method has the same form of traditional path analysis for PSE. Furthermore, under single mediator without a mediator-outcome confounder affected by exposure, it also has the same form of the results of causal mediation analysis. We also provide SAS code for settings of linear regression with exposure-mediator interaction and perform analysis in Framingham Heart Study dataset, investigating the mechanism of smoking on systolic blood pressure mediated by both cholesterol and body weight. Allowing decomposition of total effect into several rPSEs, our method contributes to investigation of complicated causal mechanisms in settings with multiple mediators.

Introduction

Mediation analysis is a technique to decompose the total effect of an exposure on an outcome into a direct effect (the effect not through a mediator) and an indirect effect (the effect through a mediator). Causal mediation analysis, defining both direct and indirect effects based on counterfactual models, extends mediation analysis to settings with nonlinearity and interaction¹⁻³. Numerous methodological techniques based on causal mediation analysis have been proposed recently, allowing different outcome scales, including additive, multiplicative, odds ratio scales, and other nonlinear models for time to event data³⁻¹¹. Most of the above techniques only consider one mediator. Under settings with multiple mediators, several approaches are available corresponding to different scientific questions. To evaluate the indirect effect mediated by all mediators, VanderWeele and Vansteelandt have proposed a regression-based method and a weighting method to estimate the direct and indirect effects mediated by all mediators at once¹². To evaluate the effect mediated by a certain combination of mediators, called path-specific effect (PSE)¹³, VanderWeele et al and Avin et al have developed methods to identify part of PSEs non-parametrically by empirical dataset^{12,13}. However, for identification of all types of PSEs, strong assumptions such as linear structural equation model (SEM), no mediator affected by other mediator, or cross-world exchangeability are required¹⁴⁻¹⁶. Assuming normal distribution for cross-world counterfactual model as well as no time-varying confounding, Daniel et al also proposed a sensitivity analysis technique to estimate the bounds of each PSE¹⁵.

Recently, alternative definitions of direct and indirect effects, i.e. randomly interventional analogues of natural direct effect (rNDE) and of natural indirect effect

(rNIE), have been used for settings with time-varying confounders ^{9,11,17}. The definition was also extended to longitudinal settings with time-varying exposures, mediators, and confounders ¹⁸⁻²¹. In this study, we extend the definitions of rNDE and rNIE to address the identification problem of PSEs in settings with multiple mediators. We first describe the notation and definitions of randomly interventional analogues of path-specific effects (rPSEs), present the non-parametric identification along with the required assumptions, and show the relation to the existing methods including path analysis and causal mediation analysis. We also provide SAS code for settings of linear regression with exposure-mediator interaction, using Framingham Heart Study dataset to investigate the mechanisms of smoking behavior on systolic blood pressure mediated by cholesterol level and weight change as example. Finally, we conclude by discussing the strengths and limitations of our method.

Notation and review of standard causal mediation analysis in a setting with two mediators

Notation and review for counterfactual models

Consider a setting with one exposure, one outcome, two mediators, two mediator-outcome confounders, and one baseline confounder as in Figure 3.1. Let A , Y , and V denote the exposure, outcome, and baseline confounders, respectively; M_1 and M_2 denote the first and second mediators, respectively. L_1 denotes the time-dependent confounders between Y and M_1 while L_2 the time-dependent confounder between Y and M_2 . Both mediator-outcome confounders (L_1 and L_2) can be affected by previous

covariates including exposure A . The causal relationship among these variables is demonstrated in Figure 3.1. Let $Y(a, m_1, m_2)$ be the counterfactual value of Y given the exposure A is set to a and the two mediators M_1 and M_2 are set to m_1 and m_2 , respectively. Let $M_2(a, m_1)$ be the counterfactual value of M_2 given A is set to a and the first mediator M_1 is set to m_1 . Let $M_1(a)$, $M_2(a)$, and $Y(a)$ be the counterfactual values of M_1 , M_2 , and Y , respectively, given A is set to a . Let G_1 and G_2 denote random draws from the distribution of the mediator M_1 and M_2 , respectively. We use similar definition for G for counterfactual models of M . For example, $G_1(a)$ is a random draw of $M_1(a)$; $G_2(a, G_1(a'))$ is a random draw of $M_2(a, G_1(a'))$, i.e. from the counterfactual value of M_2 given A is set to a and M_1 is set to $G_1(a')$. In addition, we define $G_{12}(a)$ as the random draw of $(M_1(a), M_2(a))$, i.e. counterfactual outcome of (M_1, M_2) given A is set to a . Moreover, we also make the consistency assumption^{3,5,22}. Under this assumption, $Y(a, m_1, m_2) = Y$ given $A = a$, $M_1 = m_1$, and $M_2 = m_2$. $M_2(a, m_1) = M_2$ given $A = a$ and $M_1 = m_1$. $M_1(a)$, $M_2(a)$, and $Y(a)$ are equal to M_1 , M_2 , and Y , respectively, given $A = a$.

Definitions of total effect (TE), control direct effect (CDE), natural direct effect (NDE),

natural indirect effect (NDE), and the randomly interventional analogues of these effects

Let $A = a_1$ and $A = a_0$ denote two hypothetical intervention statuses (for example, exposure and non-exposure, respectively). We use counterfactual models described above to define all effects of the exposure on the outcome by comparing two exposure levels, a_1 and a_0 . The total effect (TE) is defined as $E[Y(a_1)] - E[Y(a_0)]$. For mediation analysis, the total effect is decomposed into direct effect and indirect effect mediated by two mediators,

M_1 and M_2 . Two strategies are available for different scientific questions of interest. The first strategy assesses direct effect and indirect effect by the controlled direct effect (CDE) and the difference of TE and CDE, respectively. CDE is defined as $E[Y(a_1, m_1, m_2)] - E[Y(a_0, m_1, m_2)]$, which can be interpreted as the effect of the exposure on the outcome while two mediators, M_1 and M_2 , are intervened as certain levels, m_1 and m_2 , respectively. The difference between the total effect and the CDE can be used to estimate the extent to which the total effect blocked by setting the mediators to a certain level and is valuable for questions about policy making. For identifying CDE, we can use two assumptions:

(1) no unmeasured exposure-outcome confounding (mathematically expressed as $Y(a, m_1, m_2) \perp A | V$)

(Assumption 1)

and (2) no unmeasured mediator-outcome confounding (mathematically expressed as $Y(a, m_1, m_2) \perp M_1 | V, A, L_1$ and $Y(a, m_1, m_2) \perp M_2 | V, A, L_1, M_1, L_2$)

(Assumption 2).

Under the above assumptions, CDE can be identified as

$$Q(a_1, m_1, m_2) - Q(a_0, m_1, m_2)$$

where $Q(a, m_1, m_2) =$

$$\sum_v \{ \sum_{l_1, l_2} E[Y | v, a, l_1, m_1, l_2, m_2] \times \Pr(l_1 | v, a) \times \Pr(l_2 | v, a, l_1, m_1) \} \Pr(v)$$

$Q(a, m_1, m_2)$ is the g-formula proposed by Robins²³ while A , M_1 , and M_2 are intervened as a , m_1 , and m_2 . For any random variable W , let w denote $W = w$ in all probability function. For example, $\Pr(l_2|v, a, l_1, m_1)$ denotes $\Pr(L_2 = l_2|V = v, A = a, L_1 = l_1, M_1 = m_1)$.

For questions about investigation of causal mechanism, one often instead divides the TE into a natural direct effect (NDE) and a natural indirect effect (NIE), which are defined as the following equations^{5,12,13}:

$$NDE = \Phi(a_1, a_0) - \Phi(a_0, a_0)$$

$$NIE = \Phi(a_1, a_1) - \Phi(a_1, a_0)$$

where $\Phi(a, a')$, the standard mediation parameter, is defined as $E[Y(a, M_1(a'), M_2(a'))]$. NDE expresses the change of outcome given the exposure is changed from a_0 to a_1 , but the mediators are kept at the level they would be if the exposure is set to a_0 . In contrast, NIE expresses the change of outcome given the exposure is set to a_1 but the mediator is changed from the level it will be if exposure is set to a_0 to the level it will be if exposure is set to a_1 . To identify the standard mediation parameter (as well as NDE and NIE) non-parametrically by empirical data, we have to make four assumptions¹²:

(1) Assumption 1,

(2) no unmeasured mediator-outcome confounding (mathematically expressed as $Y(a, m_1, m_2) \perp (M_1, M_2) | V, A, L_1$)

(Assumption 2-1),

(3) no unmeasured exposure-mediator confounding (mathematically expressed as $A \perp (M_1(a), M_2(a)) | V$)

(Assumption 3),

and

(4) no mediator-outcome confounders are affected by exposure (mathematically expressed by $Y(a, m_1, m_2) \perp (M_1(a), M_2(a)) | V$)

(Assumption 4).

Although Assumption 2 and Assumption 2-1 are both interpreted as “no unmeasured mediator-outcome confounding”, the former is weaker than the latter. For example, Assumption 2-1 is violated under the presence of a M_2 - Y confounder affected by M_1 . However, Assumption 2 still holds if this confounder can be measured accurately.

Under four assumptions, $\Phi(a, a')$, NDE, and NIE can be non-parametrically identified as $Q(a, a')$, $Q(a_1, a_0) - Q(a_0, a_0)$, and $Q(a_1, a_1) - Q(a_1, a_0)$, respectively, where $Q(a, a') = \sum_v \sum_{m_1, m_2} E[Y|v, a, m_1, m_2] \Pr(m_1, m_2 | v, a') \Pr(v)$.

To ensure Assumption 4 held, L_1 and L_2 should not exist, i.e. there should be no mediator-outcome confounders affected by the exposure^{13,24}. This strong assumption can be violated even if the two mediators occurred soon after the exposure. Therefore,

alternative definitions for NDE and NIE, i.e. randomly interventional analogues of natural direct effect (rNDE) and of natural indirect effects (rNIE), have been proposed^{9,17} for settings with time-varying confounders^{18-21,25}. rNDE and rNIE are defined as follows:

$$rNDE = r\Phi(a_1, a_0) - r\Phi(a_0, a_0)$$

$$rNIE = r\Phi(a_1, a_1) - r\Phi(a_1, a_0)$$

where the $r\Phi(a, a')$, the randomly interventional analogue of $\Phi(a, a')$, is defined as $E[Y(a, G_{12}(a'))]$, which replaces $(M_1(a'), M_2(a'))$ in $\Phi(a, a')$ by $G_{12}(a')$. The rNDE expresses the change of outcome given the exposure changes from a_0 to a_1 but mediators are set to the value randomly drawn from a distribution of population with exposure is set to a_0 . The rNIE expresses the change of outcome given the exposure is set to a_1 , but the mediator changes from the value randomly drawn from the distribution in the population if the exposure were set to a_0 to the value randomly drawn from the distribution in the population if exposure were set to a_1 . The sum of rNDE and rNIE are called randomly interventional analogue of total effect (rTE).

In order to identify $r\Phi$ (as well as rNDE and rNIE), only three no unmeasured foundering assumptions (i.e., Assumption 1, Assumption 2-1, and Assumption 3) are required while Assumption 4 is not necessary anymore. Since for multiple mediators, we cannot ensure all mediators occurring right after the exposure. The mediator-outcome confounders are thus likely to be affected by exposure and Assumption 4 is violated. In

next section, we will extend the approach of randomly interventional analogue to define path-specific effects in setting with multiple mediators.

Definition and identification of randomly interventional analogues of path-specific effects

In this section, we focus on the simplest case, i.e. the setting with two mediators. The notation was introduced in previous section and the causal relationship among all covariates is shown in Figure 3.1. When there are two mediators, the number of all possible mediator combinations is four. Therefore, TE can be divided into four path-specific effects (PSEs) : (1) the path not mediated by M_1 or M_2 ($PSE_{A \rightarrow Y}$), (2) the path mediated by M_1 only ($PSE_{A \rightarrow M_1 \rightarrow Y}$), (3) the path mediated by M_2 only ($PSE_{A \rightarrow M_2 \rightarrow Y}$), and (4) the path mediated by M_1 and then by M_2 ($PSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$). In general, in settings with k mediators, the number of PSEs is 2^k , which increases exponentially. The approach is similar but become more complicated when mediator number is greater than three. Therefore, we just demonstrate our method in settings with only three mediators in Appendix 3.2 as example.

Four standard PSEs are defined as follows ^{13-15,25}:

$$PSE_{ASE} = \Phi(a_1, a_0, a_0, a_0) - \Phi(a_0, a_0, a_0, a_0)$$

$$PSE_{ASM_1 \rightarrow S} = \Phi(a_1, a_1, a_0, a_0) - \Phi(a_1, a_0, a_0, a_0)$$

$$PSE_{ASM_2 \rightarrow S} = \Phi(a_1, a_1, a_1, a_0) - \Phi(a_1, a_1, a_0, a_0)$$

$$PSE_{ASM_1 \rightarrow M_2 \rightarrow S} = \Phi(a_1, a_1, a_1, a_1) - \Phi(a_1, a_1, a_1, a_0)$$

where $\Phi(a, a', a'', a''')$ is defined as $E[Y(a, M_1(a'), M_2(a''), M_1(a'''))]$. It is worth to note that the $PSE_{A \rightarrow Y}$ is exactly the NDE and the sum of the other three PSEs is NIE. However, even under the absence of time-varying confounder, only $PSE_{A \rightarrow Y}$, $PSE_{A \rightarrow M_1 \rightarrow Y}$, and the sum of $PSE_{A \rightarrow M_2 \rightarrow Y}$ and $PSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$ can be non-parametrically identified [VW 2014, Avin 2005]. For identifying $PSE_{A \rightarrow M_2 \rightarrow Y}$ and $PSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$, besides no unmeasured confounding assumption, strong assumptions (such as linear SEM) are required^{14,15}.

Using the similar approach of rNDE and rNIE, we define randomly interventional analogues of PSEs (rPSEs): (1) $rPSE_{A \rightarrow Y}$, (2) $rPSE_{A \rightarrow M_1 \rightarrow Y}$, (3) $rPSE_{A \rightarrow M_2 \rightarrow Y}$, and (4) $rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$. Four rPSEs and rTE can be defined in terms of $r\Phi(a, a', a'', a''')$ as follows:

$$rPSE_{A \rightarrow Y} = r\Phi(a_1, a_0, a_0, a_0) - r\Phi(a_0, a_0, a_0, a_0)$$

$$rPSE_{A \rightarrow M_1 \rightarrow Y} = r\Phi(a_1, a_1, a_0, a_0) - r\Phi(a_1, a_0, a_0, a_0)$$

$$rPSE_{A \rightarrow M_2 \rightarrow Y} = r\Phi(a_1, a_1, a_1, a_0) - r\Phi(a_1, a_1, a_0, a_0)$$

$$rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y} = r\Phi(a_1, a_1, a_1, a_1) - r\Phi(a_1, a_1, a_1, a_0)$$

$$rTE = r\Phi(a_1, a_1, a_1, a_1) - r\Phi(a_0, a_0, a_0, a_0)$$

where $r\Phi(a, a', a'', a''')$ is defined as $E[Y(a, G_1(a'), G_2(a'', G_1(a''')))]$, which is the randomly interventional analogue of $\Phi(a, a', a'', a''')$. The rTE can be decomposed to four rPSEs, i.e.

$$rTE = rPSE_{A \rightarrow Y} + rPSE_{A \rightarrow M_1 \rightarrow Y} + rPSE_{A \rightarrow M_2 \rightarrow Y} + rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}.$$

Before interpreting all rPSEs and rTE, we first define five populations with hypothetical intervention on exposure (and the first mediator). Let population 1 and population 0 denote the populations with exposure set to a_1 and a_0 , respectively. Let population 1-1 and population 1-0 denote the populations with exposure set to a_1 and first mediator set to the value randomly drawn from the distribution from population 1 and population 0, respectively. Similarly, population 0-0 denotes the population with exposure set to a_0 and first mediator set to the value randomly drawn from the distribution of population 0. Then we can interpret rTE and four rPSEs based on the five populations. rTE expresses the change of outcome given the exposure changes from level a_0 to a_1 , the first mediator M_1 changes from the value randomly drawn from the distribution of population 0 to the value randomly drawn from the distribution of population 1, and M_2 changes from the value randomly drawn from the distribution of population 0-0 to the value randomly drawn from the distribution of population 1-1.; rTE captures all path from A to Y. $rPSE_{A \rightarrow Y}$ expresses the change of outcome given the exposure changes from level a_0 to a_1 but the two mediators are set to the values randomly drawn from the distribution of population 0 and population 0-0, respectively; this PSE

captures the path not mediated by M_1 or M_2 . $rPSE_{A \rightarrow M_1 \rightarrow Y}$ expresses the change of outcome given the first mediator M_1 changes from the value randomly drawn from the distribution of population 0 to the value randomly drawn from the distribution of population 1, but the exposure is set to a_0 and the second mediator M_2 is set to the value randomly drawn from a distribution of population 0-0; this PSE captures the path mediated by M_1 only without M_2 . $rPSE_{A \rightarrow M_2 \rightarrow Y}$ expresses the change of outcome given the second mediator M_2 changes from the value randomly drawn from the distribution of population 0-0 to the value randomly drawn from the distribution of population 1-0, but the exposure is set to a_1 and the first mediator M_1 is set to the value randomly drawn from a distribution of population 1; this PSE captures the path mediated by M_2 only without M_1 . Finally, $rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$ expresses the change of outcome given the second mediator M_2 changes from the value randomly drawn from the distribution of population 1-0 to the value randomly drawn from the distribution of population 1-1, but the exposure is set to a_1 and the first mediator M_1 is set to the value randomly drawn from the distribution of population 1; this PSE captures the path through M_1 followed by M_2 .

For identifying $r\Phi$ (as well as all $rPSEs$), we need to make following four no unmeasured confounding assumptions:

- (1) Assumption 1,
- (2) Assumption 2,
- (3) no unmeasured exposure-mediator confounding (mathematically expressed as $A \perp (M_1(a), M_2(a, m_1)) | V$)

(Assumption 3-1),

and (4) no unmeasured mediator-mediator confounding (mathematically expressed as $M_2(a, m_1) \perp M_1 | V, A, L_1$)

(Assumption 5).

Under the four assumptions, $r\Phi(a, a', a'', a''')$ can be non-parametrically identified as the following equations:

$$\begin{aligned}
 & r\Phi(a, a', a'', a''') \\
 &= \sum_{m_2, m_1} E[Y(a, m_1, m_2)] \Pr(M_1(a') = m_1) \sum_{m_1'} \Pr[M_2(a'', m_1') = m_2] \Pr(M_1(a''') = m_1') \\
 &= Q(a, a', a'', a''')
 \end{aligned}$$

where $Q(a, a', a'', a''') =$

$$\begin{aligned}
 & \sum_{v, m_2, m_1, l_1, l_2} E[Y|v, a, l_1, m_1, l_2, m_2] \Pr(l_1|v, a) \Pr(l_2|v, a, l_1, m_1) \Pr(m_1|v, a') \\
 & \times \sum_{m_1', l_1'} \Pr(m_2|v, a'', l_1', m_1') \Pr(l_1'|v, a'') \Pr(m_1'|a''') \Pr(v)
 \end{aligned}$$

The detail proof is provided in Appendix 3.1, proof A and B.

All types of rPSEs and rTE can be expressed in terms of Q as follows.

$$rPSE_{A \rightarrow Y} = Q(a_1, a_0, a_0, a_0) - Q(a_0, a_0, a_0, a_0)$$

$$rPSE_{A \rightarrow M_1 \rightarrow Y} = Q(a_1, a_1, a_0, a_0) - Q(a_1, a_0, a_0, a_0)$$

$$rPSE_{A \rightarrow M_2 \rightarrow Y} = Q(a_1, a_1, a_1, a_0) - Q(a_1, a_1, a_0, a_0)$$

$$rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y} = Q(a_1, a_1, a_1, a_1) - Q(a_1, a_1, a_1, a_0)$$

$$rTE = Q(a_1, a_1, a_1, a_1) - Q(a_0, a_0, a_0, a_0)$$

The definitions and identification of rPSEs in settings with three mediators are provided in Appendix 3.2.

Relation to other causal mediation analysis studies

We then discuss about the relation of our method to causal mediation analysis. For simplicity of expression, we assume that the exposure A was randomly assigned. It can be generalized to observational studies by adjusting baseline confounders V. First consider that the mediator-outcome confounders are not affected by exposure, i.e. L_1 and L_2 are empty, the randomly interventional analogue of mediation parameter $r\Phi(a, a', a'', a''')$ reduces to

$$\sum_{m_2, m_1} E[Y|a, m_1, m_2] \Pr(m_1|a') \times \sum_{m_1'} \Pr(m_2|a'', m_1') \Pr(m_1'|a''')$$

Furthermore, when $a' = a'''$, $r\Phi(a, a', a'', a')$

$$= \sum_{m_2, m_1} E[Y|a, m_1, m_2] \Pr(m_1|a') \Pr(m_2|a'', m_1)$$

The proof is provided in Appendix 3.1, proof C.

The path mediated by M_1 (the sum of $PSE_{A \rightarrow M_1 \rightarrow Y}$ and $PSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$) can be captured by $r\Phi(a_1, a_1, a_1, a_1) - r\Phi(a_1, a_0, a_1, a_0)$

$$= \sum_{m_2} \sum_{m_1} E[Y|a_1, m_1, m_2] \times [\Pr(m_1|a_1) - \Pr(m_1|a_0)] \times \Pr(m_2|a_1, m_1)$$

and $PSE_{A \rightarrow M_2 \rightarrow Y}$ can be captured by $r\Phi(a_1, a_0, a_1, a_0) - r\Phi(a_1, a_0, a_0, a_0)$

$$= \sum_{m_2} \sum_{m_1} E[Y|a_1, m_1, m_2] \times \Pr(m_1|a_0) \times [\Pr(m_2|a_1, m_1) - \Pr(m_2|a_0, m_1)].$$

Both formulas have the same form of the path-specific effects in previous studies^{13,25}.

Then consider a setting with only one mediator, i.e. the L_2 and M_2 are empty, the $r\Phi(a, a', a'', a''')$ reduces to

$$\sum_{m_1, l_1} E[Y|a, l_1, m_1] \Pr(l_1|a) \Pr(m_1|a')$$

which is the identification of $E[Y(a, G(a'))]$ ²⁵. When time-varying confounder will not be affected by exposure, i.e. all L_1 , L_2 , and M_2 are all empty, $r\Phi(a, a', a'', a''')$ reduces to

$\sum_{m_1} E[Y|a, m_1] \Pr(m_1|a')$, which is the expression of the standard mediation parameter $E[Y(a, M_1(a'))]$.

A regression based approach and illustration

In this section, we propose a regression based approach. To obtain a relatively simple close form, we aim to obtain the conditional rPSEs rather than the marginal rPSEs. We also assume no time-varying confounder affected by exposure and no mediator-mediator interaction. In order to match the notation in literature of regression based causal mediation analysis^{24,26}, we also replace the notation of baseline confounder V by C .

Consider settings with binary exposure ($A = a_1$ or $A = a_0$), continuous mediators and outcome, and no time-varying confounder affected by exposure (Figure 3.2). In addition, we also assume linear regression model allowing for exposure-mediator interactions for all continuous covariates as below:

$$E[Y|A = a, M_1 = m_1, M_2 = m_2, C = c] = \theta_0 + \theta_1 a + \theta_2 m_1 + \theta_3 m_2 + \theta_4 a m_1 + \theta_5 a m_2 + \theta_c c$$

$$E[M_2|A = a, M_1 = m_1, C = c] = \beta_0 + \beta_1 a + \beta_2 m_1 + \beta_3 a m_1 + \beta_c c$$

$$\text{and } E[M_1|A = a, C = c] = \gamma_0 + \gamma_1 a + \gamma_c c.$$

According to the formula, we can derive the following expressions for four rPSEs:

$$\text{rPSE}_{A \rightarrow Y} = \{[\theta_1 + \theta_5(\beta_0 + \beta_2\gamma_0 + \beta_2\gamma_c c + \beta_c c) + \theta_4(\gamma_0 + \gamma_c c)] + [\theta_4\gamma_1 + \theta_5(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c + \beta_2\gamma_1)]a_0 + \theta_5\beta_3\gamma_1 a_0^2\}(a_1 - a_0)$$

$$\text{rPSE}_{A \rightarrow M_1 \rightarrow Y} = \{\theta_2\gamma_1 + \theta_4\gamma_1 a_1\}(a_1 - a_0)$$

$$\text{rPSE}_{A \rightarrow M_2 \rightarrow Y} = \{\theta_3(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c) + \theta_5(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c)a_1 + \theta_3\beta_3\gamma_1 a_0 + \theta_5\beta_3\gamma_1 a_1 a_0\}(a_1 - a_0)$$

$$\text{rPSE}_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y} = \{\theta_3\beta_2\gamma_1 + [(\theta_5\beta_2\gamma_1 + \theta_3\beta_3\gamma_1)]a_1 + \theta_5\beta_3\gamma_1 a_1^2\}(a_1 - a_0)$$

The proofs are given in Appendix 3.3. Several comments merit to attention. First, we include only exposure-mediator interaction here. Similar formulas can be derived allowing for mediator-mediator interaction and even three-way interaction (interaction term among A, M₁, and M₂). In Appendix 3.3, we also show the formula including mediator-mediator interaction. Second, we propose a SAS macro for applying this formula to data in Appendix 3.4. The standard error is estimated by delta method. Third, when the baseline confounders are more than one, ie when $c = (c_1, c_2, \dots, c_p)^T$, we just

need to replace the θ_c , β_c , and γ_c by $\theta_c^T = (\theta_{c1}, \theta_{c2}, \dots, \theta_{cp})$, $\beta_c^T = (\beta_{c1}, \beta_{c2}, \dots, \beta_{cp})$, and $\gamma_c^T = (\gamma_{c1}, \gamma_{c2}, \dots, \gamma_{cp})$, respectively. Finally, under the above setting as well as no exposure-mediator interaction, $rPSE_{A \rightarrow Y}$, $rPSE_{A \rightarrow M_1 \rightarrow Y}$, $rPSE_{A \rightarrow M_2 \rightarrow Y}$, and $rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$ can be assessed by estimating θ_1 , $\alpha_1\theta_2$, $\beta_1\theta_3$, and $\alpha_1\beta_2\theta_3$, respectively, which have the same form of path analysis (or SEM) ²⁷.

Illustration

We illustrate the regression based method described above by investigating the causal mechanisms of smoking behavior on systolic blood pressure (SBP) mediated by cholesterol level and body weight. Beginning in 1948 in Framingham, Massachusetts, the original Framingham cohort consisted of 5,209 participants aged from 30 to 62 years without cardiovascular disease (CVD) history at baseline. All the participants underwent examinations at the beginning of the study and routinely every two years after that. During each exam, potential CVD risk factors were collected, including socio-demographic data, lifestyle characteristics, detailed medical history, physical examination data, and blood samples. Further details on the design of FHS are described elsewhere [26, 37]. Four exclusion criteria are listed below: (1) death or loss to follow up during the period before exam 7 (the end of follow-up); (2) no record at baseline on weight, height, smoking status, former smoking history, SBP, or total cholesterol; (3) diagnosis of diabetes, cancer, or CVD at baseline; and (4) value for smoking status or BMI missing more than once. In addition, we also eliminate those who quit smoke in order to focus on the current smoker versus non-smoker comparison. After these exclusions, 2,993 participants are eligible for analysis. SBP (mm-Hg) at exam 7 is the

outcome Y and smoking amount is exposure of interest (comparing smoking for 30 cigarettes per day vs. nonsmoking). The cholesterol level (mg/dL) at exam 4 and BMI (kg/m^2) at exam 6 are two mediators M_1 and M_2 . We include gender, age (years), and baseline smoking status (smoker vs. non-smoker) as our baseline confounders. A linear regression model is fit for SBP on the cholesterol, BMI, smoking, the interaction between smoking and cholesterol level, the interaction between smoking and BMI, and the baseline covariates (gender and age). A linear regression model for BMI is fit on the smoking, cholesterol level, their interaction, and baseline covariates. A linear regression model for cholesterol level is fit on the smoking and baseline covariates. Confidence intervals are obtained using the delta method. The SAS code in the context of the rPSE decomposition is provided in the Appendix 3.4. We use this decomposition and these methods so that we can separate the effect of smoking on SBP mediated directly through cholesterol to SBP versus that which changes BMI through changing cholesterol.

Results are summarized in Table 3.1. Of these 4 rPSEs, the direct effect and effect mediated by cholesterol level only are dominant highlighting the important role of cholesterol in this context. The effect mediated by BMI is non-significant but has different direction from total effect, might indicating the possibility that the adverse effect of smoking on increasing blood pressure is partially concealed by body weight loss. However, this path is balanced by another path via cholesterol level and then BMI. Our results provide evidence that the cholesterol level might play an important role in the mechanism of smoking on SBP.

Discussion

To understand the mechanisms with multiple mediators, it is necessary to investigate all PSEs. When the mediators affect to each other, the analysis will become complicated since one path might be shared by several PSEs. In the setting with two mediators, the path from A to M_1 belongs to both $PSE_{A \rightarrow M_1 \rightarrow Y}$ and $PSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$. Therefore, based on standard causal mediation analysis, $PSE_{A \rightarrow M_1 \rightarrow Y}$ and $PSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$ cannot be identified separately. Traditional methods such as SEM or path analysis²⁷, however, deal with the identification problem by making strong assumption such as linear regression model for all covariates without interaction, which may not be applicable in more complicated settings. In addition, a linear outcome model is not feasible for most epidemiologic studies in which outcome scales are usually binary or time-to-event. Daniel et al¹⁵ has proposed a method to identify and estimate four PSEs. This method requires four assumptions: (1) consistency, (2) no unmeasured confounding among exposure, mediators, and outcome, (3) no mediator-outcome confounders affected by exposure, and (4) normal distribution for the cross-world counterfactual value of the first mediator. The first two assumptions are shared by all causal inference methods and the third one, shared by almost all causal mediation analyses, is a relatively strong assumption which still can be violated even the two mediators occurs after exposure immediately. Under the four assumptions, Daniel et al proposed a sensitivity analysis technique to estimate the bounds of the four PSEs¹⁵. Our approach, based on definition of randomly interventional analogue, requires only the first two assumptions described above to identify rPSEs non-parametrically. It allows for flexible choices for all types of scales and regression models. In addition, it also adjusts for time-varying confounders,

which is applicable for broader settings. However, as discussed above, the definition of the effects are slightly different.

Several limitations concerning our method merit discussion. First, rather than the standard definitions based on cross-world counterfactual outcomes, our method used definitions of randomly interventional analogue, which is not exactly how the mechanisms perform in nature. In addition, the sum of all rPSEs is rTE, rather than total effect of exposure on outcome. However, these effects can be examined in principle in randomized controlled trial, while PSEs cannot since PSEs are defined using cross-world counterfactual outcomes²⁴. Second, assumptions of no unmeasured confounding are required for accurate rPSE estimates. To ensure these assumptions held, researchers should collect all potential confounders as comprehensively as possible. When collection of all covariates is impossible, sensitivity analysis techniques could be developed to assess the extent of bias due to assumption violation. Finally, our SAS code only allows linear regression model with exposure-mediator interaction. Methods allowing binary or time to event outcome could be developed but are not yet available. For applying this method more broadly, more methods and corresponding software should be developed in the future.

In conclusion, our study provides a framework to decompose rTE into several rPSEs mediated by all possible combinations of mediators, extending the standard analysis method to settings with interaction, non-linearity, and time-varying confounders affected by exposure. Our method contributes to investigation of complicated causal mechanisms in settings with multiple mediators.

Reference

1. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 1992;143-155.
2. Pearl J. Direct and indirect effects. Proceedings of the seventeenth conference on uncertainty in artificial intelligence: Morgan Kaufmann Publishers Inc., 2001;411-420.
3. VanderWeele T, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface* 2009;2:457-468.
4. van der Laan MJ, Petersen ML. Direct effect models. *The international journal of biostatistics* 2008;4(1).
5. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology* 2009;20(1):18-26.
6. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychological methods* 2010;15(4):309.
7. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology (Cambridge, Mass.)* 2010;21(4):540.
8. VanderWeele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *American journal of epidemiology* 2010;172(12):1339-1348.
9. Didelez V, Dawid P, Geneletti S. Direct and indirect effects of sequential treatments. *arXiv preprint arXiv:1206.6840* 2012.
10. Tchetgen EJT, Shpitser I. Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis. *The Annals of Statistics* 2012;40(3):1816-1845.

11. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological methods* 2013;**18**(2):137.
12. VanderWeele TJ, Vansteelandt S. Mediation Analysis with Multiple Mediators. *Epidemiol Method* 2014;**2**(1):95-115.
13. Avin C, Shpitser I, Pearl J. Identifiability of path-specific effects. *Department of Statistics, UCLA* 2005.
14. Albert JM, Nelson S. Generalized causal mediation analysis. *Biometrics* 2011;**67**(3):1028-1038.
15. Daniel R, De Stavola B, Cousens S, Vansteelandt S. Causal mediation analysis with multiple mediators. *Biometrics* 2015;**71**(1):1-14.
16. Taguri M, Featherstone J, Cheng J. Causal mediation analysis with multiple causally non-ordered mediators. *Statistical methods in medical research* 2015:0962280215615899.
17. Geneletti S. Identifying direct and indirect effects in a non-counterfactual framework. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2007;**69**(2):199-215.
18. Zheng W, van der Laan MJ. Causal mediation in a survival setting with time-dependent mediators. 2012.
19. VanderWeele TJ, Tchetgen Tchetgen E. Mediation Analysis with Time-Varying Exposures and Mediators. 2014.
20. Lin S-H, Young J, Tchetgen Tchetgen E, VanderWeele TJ. Parametric mediational g-formula approach to mediation analysis with time-varying exposures, mediators

- and confounders: an application to smoking, weight, and blood pressure.
Epidemiology (submitting) 2015.
21. Lin S-H, Tchetgen Tchetgen E, VanderWeele TJ. Mediation analysis with survival outcome and time-varying exposures, mediators, and confounders: a case study of the Framingham Heart Study. *submitting* 2015.
 22. Pearl J. Causal inference in statistics: An overview. *Statistics Surveys* 2009;**3**:96-146.
 23. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling* 1986;**7**(9):1393-1512.
 24. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction* Oxford University Press, 2015.
 25. VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology* 2014;**25**(2):300-306.
 26. VanderWeele TJ. Concerning the consistency assumption in causal inference. *Epidemiology* 2009;**20**(6):880-883.
 27. MacKinnon DP. *Introduction to statistical mediation analysis* Routledge, 2008.

Figure 3.1. Causal diagrams for a setting with two mediators and time-varying mediator-outcome confounders affected by exposure.

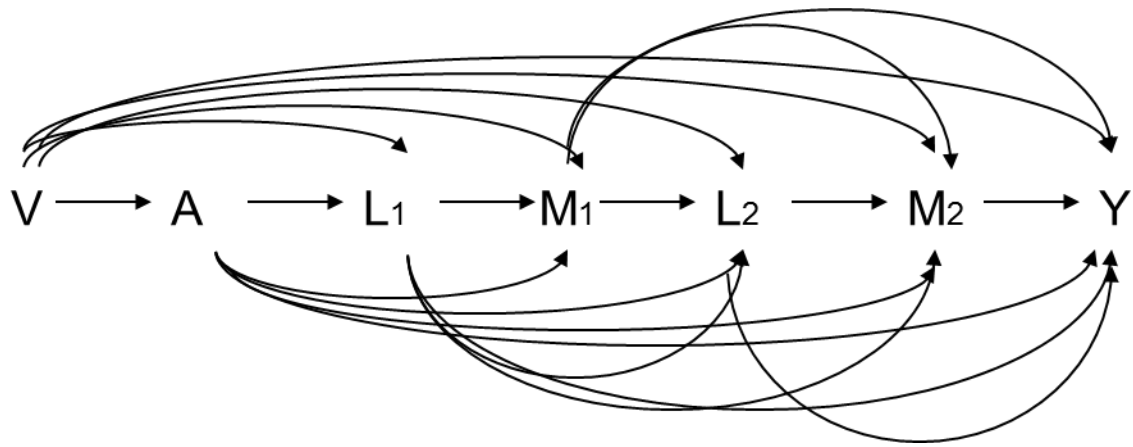


Figure 3.2. Causal diagrams for a setting with two mediators but no time-varying mediator-outcome confounders.

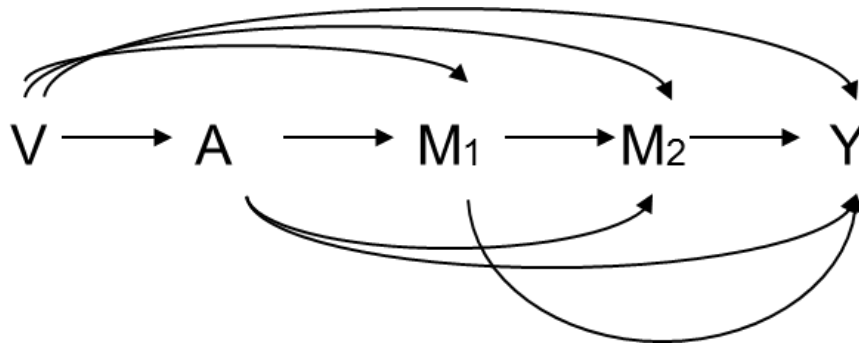


Table 3.1. Proportions of the effect of smoking on systolic blood pressure mediated by cholesterol and/or body mass index.

	Effects	(95 % CI)	p-value	Proportion	(95 % CI)	p-value
	Attributable					
$rPSE_{A \rightarrow Y}$	1.9560	(-0.6758, 4.5879)	0.1452	0.6464	(0.2704, 1.0223)	0.0008
$rPSE_{A \rightarrow M_1 \rightarrow Y}$	1.0348	(0.4009, 1.6687)	0.0014	0.3419	(0.0109, 0.6730)	0.0429
$rPSE_{A \rightarrow M_2 \rightarrow Y}$	-0.1655	(-0.8291, 0.4980)	0.6248	-0.05470	(-0.2893, 0.1799)	0.6476
$rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y}$	0.2009	(0.0369, 0.3650)	0.0164	0.06640	(-0.0152, 0.1480)	0.1108
Effect via M1 (with/without M2)	1.2357	(0.5488, 1.9226)	0.0004	0.4083	(0.0173, 0.7994)	0.0407
Effect via M2 (with/without M1)	0.03541	(-0.6426, 0.7134)	0.9184	0.01170	(-0.2105, 0.2339)	0.9178
Effect via M1 or M2	1.0702	(0.1195, 2.0209)	0.0274	0.3536	(-0.0223, 0.7296)	0.0652
rTE	3.0262	(0.2679, 5.7846)	0.0315	1		

Supplementary Files

Appendix 1: Description of mgformula SAS macro for parametric mediational g-formula

The implementation of causal mediation analysis with time-varying exposures, mediators, and confounders

Introduction

The mgformula SAS macro is designed to conduct causal mediation analysis with time-varying exposures, mediators, and confounders in longitudinal data. The outcome can be a continuous, binary, or time-to-event (i.e. survival) variable. The macro provides point estimates and confidence intervals for the overall effect, the direct effect and the indirect effect. The estimates are unbiased, under the model and identifiability assumptions, which are described elsewhere¹⁻³. Because the mgformula macro is specifically developed for time-varying mediators, this macro allows for a fixed exposure but not for a fixed mediator. For settings with a fixed exposure and mediator, the *mediation* macro is appropriate and can be freely downloaded on the website (<http://www.hsph.harvard.edu/causal/software/>).

Notation and review for mediational g-formula

Consider exposures, mediators, and confounders that vary over time in a longitudinal data setting with measurements at time 0, 1, 2, ..., T-1. Let (L(0), L(1), ..., L(T-1)), (M(0), M(1), ..., M(T-1)), and (A(0), A(1), ..., A(T-1)) denote values of the time-varying confounders, mediators, and exposures at periods 0, 1, ..., T-1, with the final outcome of interest Y (or S(T-1) for a survival outcome). The causal relationships among

these variables are given in Figure S.1. To accommodate compatibility with the regular g-formula macro, although the most common causal relationships follow the order of L, A, and M, the current mgformula macro assumes the ordering of the variables as L, M, and A. Thus, the data should be structured so that the mediator from the current visit is entered along with the exposure from the subsequent visit.

For any variable W , let $\overline{W(t)} = (W(0), W(1), \dots, W(t))$ and let $\overline{W} = \overline{W(T-1)} = (W(0), W(1), \dots, W(T-1))$. Let $Y_{\overline{am}}$ be the counterfactual outcome given \overline{A} is set to \overline{a} and if \overline{M} is set to \overline{m} . Let $M_{\overline{a}}(t)$ be the counterfactual value of $M(t)$ given \overline{A} is set to \overline{a} . Let $G_{\overline{a}}(t)$ denote a random draw from the distribution of the mediator $M(t)_{\overline{a}}$. In this setting, the randomly interventional analogues of total effect (rTE), of natural direct effect (rNDE), and of natural indirect effect (rNIE) (representing the overall effect, the direct effect, and the indirect effect, respectively) are defined as $E[Y_{\overline{aG\overline{a}}}] - E[Y_{\overline{a^*G\overline{a^*}}}]$, $E[Y_{\overline{aG\overline{a^*}}}] - E[Y_{\overline{a^*G\overline{a^*}}}]$, and $E[Y_{\overline{aG\overline{a}}}] - E[Y_{\overline{aG\overline{a^*}}}]$, respectively. The rTE can be decomposed into the rNDE and rNIE, i.e. $E[Y_{\overline{aG\overline{a}}}] - E[Y_{\overline{a^*G\overline{a^*}}}] = (E[Y_{\overline{aG\overline{a^*}}}] - E[Y_{\overline{a^*G\overline{a^*}}}]) + (E[Y_{\overline{aG\overline{a}}}] - E[Y_{\overline{aG\overline{a^*}}}]$). To identify rTE, rNDE, and rNIE, we need to make the three following assumptions for all t : (1) $Y_{\overline{am}} \perp A(t) \mid \overline{A(t-1)}, \overline{M(t)}, \overline{L(t)}$ (no exposure-outcome confounding conditional on the past variables, while \perp denotes independent); (2) $Y_{\overline{am}} \perp M(t) \mid \overline{A(t-1)}, \overline{M(t-1)}, \overline{L(t)}$ (no mediator-outcome confounding conditional on the past variables); and (3) $M_{\overline{a}}(t) \perp A(t) \mid \overline{A(t-1)}, \overline{M(t-1)}, \overline{L(t)}$ (no exposure-mediator confounding conditional on the past variables)¹. Under the three assumptions, the rTE, rNDE and rNIE are identified non-parametrically by the following equations:

$$rTE = Q(\bar{a}, \bar{a}) - Q(\bar{a}^*, \bar{a}^*)$$

$$rNDE = Q(\bar{a}, \bar{a}^*) - Q(\bar{a}^*, \bar{a}^*)$$

$$rNIE = Q(\bar{a}, \bar{a}) - Q(\bar{a}, \bar{a}^*)$$

where $Q(\bar{a}_1, \bar{a}_2)$

$$= \sum_{\bar{m}} \sum_{\bar{l}} E[Y | \bar{a}_1, \bar{m}, \bar{l}] \prod_{t=0}^{T-1} P(l(t) | \bar{a}_1(t-1), \bar{m}(t-1), \bar{l}(t-1)) \times \\ \sum_{\bar{l}'} \prod_{t=0}^{T-1} P(m(t) | \bar{a}_2(t-1), \bar{m}(t-1), \bar{l}'(t)) P(l'(t) | \bar{a}_2(t-1), \bar{m}(t-1), \bar{l}'(t-1))$$

We refer to $E[Y_{\bar{a}G\bar{a}^*}]$ as the mediation parameter and $Q(\bar{a}_1, \bar{a}_2)$ as the mediational g-formula. Using the mgformula macro, we can derive the $Q(\bar{a}_1, \bar{a}_2)$ parametrically and then calculate the point estimates and confidence intervals for the rTE, rNDE, and rNIE.

Quick start introduction

You can follow the four steps below to implement the mgformula macro to your data within 5 minutes. The details are described in the next section (“Basic SAS Macro and corresponding parameters”).

Step 1. Prepare your dataset: Prepare a dataset in a person-time format to include a subject id, time of follow-up (t), the outcome (Y), the exposures (A), the mediators (M), and the confounders (L), as the format in Table S.1. For any random variable W (W can

be Y, A, M, or L), $W_i(t)$ represents the observation of W for individual i at time t. Make sure the data is completed (no censoring or competing risk).

Step 2. Save and call mgformula macro: Save mgformula macroscript ("mgformula.sas") in a new file ("file_direction"), and at the top of this program use the following %include statement :

```
%include 'file_direction/mgformula.sas';
```

Step 3. Define a local macro variable (cov): Specify a gformula macro variable, cov. An example of the required syntax for the simple case of one time-varying confounder, L, and total times of follow up (T) = 5 is provided below. Time-varying covariates must be listed in the order of L, M, and A:

```
%let cov=
```

```
ncov=3 ,
```

```
cov1= L, cov1otype=1 , cov1ptype=conbin,
```

```
cov2= M, cov2otype=1 ,cov2lptype=conbin,
```

```
cov3= A, cov3otype=1, cov3ptype=conbin,
```

```
seed= 9458;
```

Step 4. Specify mgformula parameters and run: run mgformula macro by the following statement:

%mgformula(datain= , id = , time= , timepoints=, yvar= , yreg= ,avar= , mvar= , a1value= , a0value= , nboot=) ;

Input the name of the data set (datain=), the subject id (id=), the time of observation (time=), the overall follow-up time (timepoints=), the outcome variable (yvar=) and its model (yreg=), the exposure variable (avar=), and the mediator variable (mvar=). Then specify the exposure level a (a1value=) and the baseline level of the exposure a* (a0value=), and the number of bootstrap samples (nboot=). Three choices for the outcome variable models are available: continuous variable (yreg=conteofu), binary variable (yreg=bineofu), and binary survival variable (yreg=binsurv).

Basic SAS Macro and corresponding parameters

In order to activate the mgformula macro, four steps are required.

1. Prepare your dataset

The dataset has to include the following variables: the subject identification (id), time of follow up (t), the outcome (Y), the exposures (A), the mediators (M), and the confounders (L(s)). For simplicity, we assume no baseline confounder and only one time-varying confounder (L) for the following description. If lagged values of the variable (var) are to be used as predictors, then var_l1, var_l2 and var_l3 should be included in the dataset. Here, var_t denotes the variable value measured at t earlier time point (t = 1, 2, or 3). We will describe the lag model in covXptype parameter section and Table S.3. The data must be arranged with one record per subject per time point (Table

S.1-1 and Table S.1-2). The outcome in mgformula macro allows a binary, continuous, or survival variable. When analyzing binary or continuous outcome (Y), the Y_l1 is recommended to be included in the data as another time-varying confounder. Please refer to step 3 “define cov parameter” for more details about defining covariates. The mgformula macro internally involves several calls to the “gformula macro”⁴ for parametric estimation of the standard gformula and relies on the same general input data structure. Additional details regarding input dataset requirements are mentioned in the gformula SAS macro instruction⁴. At this time, this approach does not handle competing risks or censoring, which will be part of our future work. Therefore, the data must be restricted to observations with complete follow-up and no competing risk events.

2. Save and call mgformula macro

Open a new SAS session, input the data, save mgformula macroscript ("mgformula.sas") in a new file ("file_direction"), and at the top of this program use the following %include statement:

```
%include 'file_direction/mgformula.sas';
```

3. Define cov parameter

In order to specify the baseline confounders and the time-varying covariates (including exposures, mediators, and confounders), and their models, we define the "cov" parameters by the following statement:

```
%let cov=

    fixedcov = ,

    ncov=3 ,

    cov1= L, cov1otype=1 , cov1ptype=conbin,

    cov2= M, cov2otype=1 ,cov2lptype=conbin,

    cov3= A, cov3otype=1, cov3ptype=conbin,

    seed= 9458;
```

The *cov* parameters consist of several parameters including *fixedcov*, *ncov*, *covX*, *covXotype*, *covXptype*, etc, where X = 1, 2, and 3. The limit of X is up to 30. The *fixedcov* parameter specifies baseline confounders (and their models) and the other parameters specify the time-varying covariates (and their models). All time-varying exposures, mediators, and confounders should be included in time-varying covariates in this step, and will be further distinguished in the next step. We describe each parameter as follows.

fixedcov (optional)

The *fixedcov* parameter specifies the baseline confounder(s). This parameter is optional, so it is allowed to be empty. All covariates in the *fixedcov* parameter will be included as predictors of all time-varying covariates models and the outcome model. The relationship between baseline confounder(s) and time-varying covariates (and the outcome) are assumed linear. The macro cannot transform variables automatically so we have to do it manually if we want to. For example, if we plan to include variable "age" as a baseline confounder and fit a quadratic linear model, we should create an additional variable "age2" which is equal to age^2 , and specify this parameter as "*fixedcov* = *age age2*".

The time-varying covariates and their models are defined by several parameters listed below.

ncov (required)

The *ncov* parameter indicates the number of time-varying covariates. For example, if we specify three time-varying covariates: time-varying confounder (L), mediator (M), and exposure (A), then we should specify this parameter as "*ncov*=3".

For all $X = 1, \dots, \text{ncov}$:

covX (required for $X = 1$)

The *covX* parameter specifies the X^{th} time-varying covariate to be included in the model. The users should specify the mediator and the exposure as the last second covariate and the last covariate, respectively. In the setting with three time-varying covariates, we should specify *cov1*, *cov2*, and *cov3* as the time-varying confounder (L), the mediator (M), and the exposures (A), respectively. For continuous or binary outcome, we recommend to include *y_l1* as one of the time-varying confounders (we should create *y_l1* when preparing the dataset).

covXotype (required when covX not empty)

covXptype (required when covX not empty)

The *mgformula* macro fits a parametric model for the distributions of outcome and all time-varying covariates conditional on previous covariates. Thus, each time-varying covariate is necessary to specify both how it is treated as the dependent variable (outcome type or "otype") and how it is treated as a predictor for the outcome and all later covariate density models (predictor type or "ptype"). For example, given *cov1* is time-varying confounder, L. *L(1)* is used to predict the future variables such as *M(1)*, *A(1)*, *L(2)*, ... , and *Y*. On the other hand, the variables previous to *L(1)*, such as *L(0)*, *M(0)*, and *A(0)* (as well as baseline confounders if they exist), is used as predictors for the distribution of *L(1)*. For the model with *covX* as predictor, we specify *covXptype* parameter to indicate a certain "predictor type" for *covX*. For the model with *covX* as the dependent variable, we specify *covXotype* parameter to indicate a certain "outcome type" for *covX*. Note that at this time, the macro can allow a covariate with only one otype and one ptype for different models. The complete choices for ptype and otype are summarized in Table 2 and Table

3 of the gformula SAS macro instruction⁴. Here, we introduce the most common types as follows.

Three common outcome types (*covXotype* = 1, 2, and 3) are summarized in Table S.2. When *covX* is a binary variable, specify *covXotype*=1. When *covX* is a survival variable, specify *covXotype*=2. When *covX* is a continuous variable, specify *covXotype*=3.

Several common predictor types for time-varying covariates into the models are summarized in Table S.3. The most common choice is "covXptype = conbin", where "con" indicates concurrent model and "bin" the binary or linear model. The "con" can be replaced by lag1, lag2, or lag3 for different lag models. Again, if lagged values of the variable (var) are used as predictors, then var_l1, var_l2 and var_l3 should be prepared and included in the dataset. In addition, the bin can also be replaced by qdc, cub, or zqdc, for the quadratic, cubic, or zero quadratic model, respectively. For example, the cubic lag2 model for cov1 (given cov1 = L) can be specified as "cov1ptype=lag2cub" and both L_l1 and L_l2 should be created in the dataset. All possible types are precisely described in the gformula macro instruction (GFORMULA SAS macro instruction, Table 3 in the Models section).

The transformations of the covariates required for predictors are generated automatically by them GFORMULA macro and do not need to be included in the dataset. Also, to model the full joint distribution of the covariates and outcome, the macro imposes the order of covariates as listed and uses the concurrent values of earlier covariates to predict all subsequent covariates.

seed (optional)

Finally, the **seed** parameter specifies the random numbers seed. The default number is 7834.

The cov parameters are the same as the gformula macro from "fixedcov" to "seed". Please refer to the gformula instruction for detailed description.

4. Specify mgformula parameter and run

Finally, we specify the mgformula parameters (datain, id, time, timepoints, yvar, yreg, avar, mvar, alvalue, a0value, and nboot) and call the mgformula macro by the statement:

```
%mgformula(datain= , id = , time= , timepoints=, yvar= , yreg= ,avar= , mvar= ,  
alvalue= , a0value= , nboot=) ;
```

We first input the name of the dataset (datain=), the identification (id=), the time of observation (time=), the overall follow-up time (timepoints=), the outcome variable (yvar=) and its model (yreg=), the exposure variable (avar=), and the mediator variable

(mvar=). Then we specify the exposure level a (a1value=) and the baseline level of the exposure a* (i.e. non-exposure level) (a0value=), and the number of bootstrapping samples (nboot=). Three choices for the outcome variable models are available: continuous variable (yreg=conteofu), binary variable (yreg=bineofu), and binary survival variable (yreg=binsurv). The paramters datain, yvar, and yreg correspond to the GFORMULA parameters data, outc, and outctype, respectively.

Output

The output file is like Table S.4. The estimation, standard error, and 95% confidence interval for the total effect (TE), direct effect (DE), and indirect effect (IE) are generated. We use the original output of Table 1.4 in the manuscript. The regression models and the corresponding code we used for the analysis are provided as follows.

Regression models:

Model 1:

$$\begin{aligned}
 E[Y|x_1(4), x_2(4), x_3(4), x_4(4), x_5(4), x_1(3), x_2(3), x_3(3), x_4(3), x_5(3), c] = & \beta(1)_0 + \beta(1)_{11}x_1(4) + \\
 & \beta(1)_{12}(x_1(4))^2 + \beta(1)_{11p}x_1(3) + \beta(1)_{12p}(x_1(3))^2 + \beta(1)_{21}I(x_2(4)=1) + \beta(1)_{22}I(x_2(4)=2) + \\
 & \beta(1)_{21p}I(x_2(3)=1) + \beta(1)_{22p}I(x_2(3)=2) + \beta(1)_{31}x_3(4) + \beta(1)_{32}(x_3(4))^2 + \beta(1)_{31p}x_3(3) + \\
 & \beta(1)_{32p}(x_3(3))^2 + \beta(1)_{41}x_4(4) + \beta(1)_{42}(x_4(4))^2 + \beta(1)_{41p}x_4(3) + \beta(1)_{42p}(x_4(3))^2 + \beta(1)_{51}x_5(4) \\
 & + \beta(1)_{52}(x_5(4))^2 + \beta(1)_{51p}x_5(3) + \beta(1)_{52p}(x_5(3))^2 + \beta(1)_c c
 \end{aligned}$$

Model 2:

$$\begin{aligned} \text{logit}(E[I(X_5(t)=0)|x_1(t),x_2(t),x_3(t),x_4(t),x_1(t-1),x_2(t-1),x_3(t-1),x_4(t-1),x_5(t-1),c]) &= \beta(2)_0 + \\ \beta(2)_{11}x_1(t) + \beta(2)_{12}(x_1(t))^2 + \beta(2)_{11p}x_1(t-1) + \beta(2)_{12p}(x_1(t-1))^2 + \beta(2)_{21}I(x_2(t)=1) + \\ \beta(2)_{22}I(x_2(t)=2) + \beta(2)_{21p}I(x_2(t-1)=1) + \beta(2)_{22p}I(x_2(t-1)=2) + \beta(2)_{31}x_3(t) + \beta(2)_{32}(x_3(t))^2 + \\ \beta(2)_{31p}x_3(t-1) + \beta(2)_{32p}(x_3(t-1))^2 + \beta(2)_{41}x_4(t) + \beta(2)_{42}(x_4(t))^2 + \beta(2)_{41p}x_4(t-1) + \\ \beta(2)_{42p}(x_4(t-1))^2 + \beta(2)_{51p}x_5(t-1) + \beta(2)_{52p}(x_5(t-1))^2 + \beta(2)_{c0} \end{aligned}$$

Model 3:

$$\begin{aligned} E[\log(X_5(t))|x_1(t),x_2(t),x_3(t),x_4(t), x_5(t)>0, x_1(t-1), x_2(t-1), x_3(t-1), x_4(t-1), x_5(t-1),c] &= \beta(3)_0 \\ + \beta(3)_{11}x_1(t) + \beta(3)_{12}(x_1(t))^2 + \beta(3)_{11p}x_1(t-1) + \beta(3)_{12p}(x_1(t-1))^2 + \beta(3)_{21}I(x_2(t)=1) + \\ \beta(3)_{22}I(x_2(t)=2) + \beta(3)_{21p}I(x_2(t-1)=1) + \beta(3)_{22p}I(x_2(t-1)=2) + \beta(3)_{31}x_3(t) + \beta(3)_{32}(x_3(t))^2 + \\ \beta(3)_{31p}x_3(t-1) + \beta(3)_{32p}(x_3(t-1))^2 + \beta(3)_{41}x_4(t) + \beta(3)_{42}(x_4(t))^2 + \beta(3)_{41p}x_4(t-1) + \\ \beta(3)_{42p}(x_4(t-1))^2 + \beta(3)_{51p}x_5(t-1) + \beta(3)_{52p}(x_5(t-1))^2 + \beta(3)_{c0} \end{aligned}$$

Model 4:

$$\begin{aligned} E[X_4(t)|x_1(t),x_2(t),x_3(t), x_1(t-1), x_2(t-1), x_3(t-1), x_4(t-1), x_5(t-1),c] &= \beta(4)_0 + \beta(4)_{11}x_1(t) + \\ \beta(4)_{12}(x_1(t))^2 + \beta(4)_{11p}x_1(t-1) + \beta(4)_{12p}(x_1(t-1))^2 + \beta(4)_{21}I(x_2(t)=1) + \beta(4)_{22}I(x_2(t)=2) + \\ \beta(4)_{21p}I(x_2(t-1)=1) + \beta(4)_{22p}I(x_2(t-1)=2) + \beta(4)_{31}x_3(t) + \beta(4)_{32}(x_3(t))^2 + \beta(4)_{31p}x_3(t-1) + \end{aligned}$$

$$\beta(4)_{32p}(x_3(t-1))^2 + \beta(4)_{41}x_4(t) + \beta(4)_{42}(x_4(t))^2 + \beta(4)_{41p}x_4(t-1) + \beta(4)_{42p}(x_4(t-1))^2 + \\ \beta(4)_{51p}x_5(t-1) + \beta(4)_{52p}(x_5(t-1))^2 + \beta(4)_{c}c$$

Model 5:

$$E[X_3(t)|x_1(t), x_2(t), x_1(t-1), x_2(t-1), x_3(t-1), x_4(t-1), x_5(t-1), c] = \beta(5)_0 + \beta(5)_{11}x_1(t) + \\ \beta(5)_{12}(x_1(t))^2 + \beta(5)_{11p}x_1(t-1) + \beta(5)_{12p}(x_1(t-1))^2 + \beta(5)_{21}I(x_2(t)=1) + \beta(5)_{22}I(x_2(t)=2) + \\ \beta(5)_{21p}I(x_2(t-1)=1) + \beta(5)_{22p}I(x_2(t-1)=2) + \beta(5)_{31p}x_3(t-1) + \beta(5)_{32p}(x_3(t-1))^2 + \beta(5)_{41}x_4(t) + \\ \beta(5)_{42}(x_4(t))^2 + \beta(5)_{41p}x_4(t-1) + \beta(5)_{42p}(x_4(t-1))^2 + \beta(5)_{51p}x_5(t-1) + \beta(5)_{52p}(x_5(t-1))^2 + \beta(5)_{c}c$$

Model 6:

$$E[X_2(t)|x_1(t), x_1(t-1), x_2(t-1), x_3(t-1), x_4(t-1), x_5(t-1), c] = \beta(6)_0 + \beta(6)_{11}x_1(t) + \beta(6)_{12}(x_1(t))^2 \\ + \beta(6)_{11p}x_1(t-1) + \beta(6)_{12p}(x_1(t-1))^2 + \beta(6)_{21p}I(x_2(t-1)=1) + \beta(6)_{22p}I(x_2(t-1)=2) + \\ \beta(6)_{31p}x_3(t-1) + \beta(6)_{32p}(x_3(t-1))^2 + \beta(6)_{41}x_4(t) + \beta(6)_{42}(x_4(t))^2 + \beta(6)_{41p}x_4(t-1) + \\ \beta(6)_{42p}(x_4(t-1))^2 + \beta(6)_{51p}x_5(t-1) + \beta(6)_{52p}(x_5(t-1))^2 + \beta(6)_{c}c$$

Model 7:

$$E[X_1(t)|x_1(t-1), x_2(t-1), x_3(t-1), x_4(t-1), x_5(t-1), c] = \beta(7)_0 + \beta(7)_{11p}x_1(t-1) + \beta(7)_{12p}(x_1(t-1))^2 \\ + \beta(7)_{21p}I(x_2(t-1)=1) + \beta(7)_{22p}I(x_2(t-1)=2) + \beta(7)_{31p}x_3(t-1) + \beta(7)_{32p}(x_3(t-1))^2 + \beta(7)_{41}x_4(t) \\ + \beta(7)_{42}(x_4(t))^2 + \beta(7)_{41p}x_4(t-1) + \beta(7)_{42p}(x_4(t-1))^2 + \beta(7)_{51p}x_5(t-1) + \beta(7)_{52p}(x_5(t-1))^2 + \\ \beta(7)_{c}c$$

Where c denotes the combination of all fixed covariates; Y denotes the outcome variable; $x_i(t)$ ($i = 1, 2, \dots, 5$) denotes the covi measured at time t ($t=0,1,\dots,4$); $\beta(j)$ denotes the coefficients at model j ($j = 1, 2, \dots, 7$).

Code:

```
%let mycov=

    fixedcov = sex age_0 age2_0 height_0 height2_0 edu2 edu3 edu4 edu5 edu6
    edu7 work2 work3 work4 work5 work6 work7 btuser1 btuser2 marital1 marital2,

    ncov=5,

    cov1= sysbp_l1, cov1otype=3, cov1ptype=qdc, cov1wherem=exam ne
    0,

    cov2 = antihbp, cov2otype=3, cov2ptype= skpcat, cov2skip=0 2,
    cov2knots= 1 2, cov2wherem= exam ne 0,

    cov3= chol, cov3otype=3, cov3ptype=skpqdc, cov3wherem= exam ne 0,

    cov4= bmi, cov4otype=3, cov4ptype= lag1qdc, cov4wherem=exam ne 0,

    cov5= cig, cov5otype=4, cov5ptype= skpqdc, cov5skip = 3, cov5wherem=
    exam ne 0, cov5class=cig3,

    seed= 7834;
```

```
%mgformula(datain=data_fhs , id =pid , time=exam , timepoints=5,
yvar=sysbp_ , yreg=conteofu, avar=cig , mvar=bmi, alvalue=20, a0value=0, nboot
= 2) ;
```

Acknowledgement:

The authors are grateful to Priyanka Jain for the insightful discussions for the mgformula SAS macro.

Reference

1. VanderWeele TJ, Tchetgen Tchetgen E. Mediation Analysis with Time-Varying Exposures and Mediators. 2014.
2. Sheng-Hsuan Lin JY, Eric TchetgenTchetgen, Miguel Hernán, Tyler J. VanderWeele. Parametric mediational g-formula approach to mediation analysis with time-varying exposures, mediators and confounders: an application to smoking, weight, and blood pressure. *Epidemiology (submitting)* 2015.
3. Sheng-Hsuan Lin EJTT, Tyler J. VanderWeele. Mediation analysis with survival outcome and time-varying exposures, mediators, and confounders: a case study of the Framingham Heart Study. *submitting* 2015.
4. Roger Logan ST, Jessica Young, Sally Picciotto, Miguel A. Hernán. GFORMULA

SAS MACRO - Estimates the mean of a dichotomous outcome at end of follow-up under general interventions on time-varying treatments in observational studies using the parametric g-formula.

<http://www.hsph.harvard.edu/causal/software/> 2015.

Table S.1-1. General form of person-time dataset when outcome is continuous or binary variable measured at the end of follow up (Y_i).

id	t	Y_i	$A_i(t)$	$M_i(t)$	$L_i(t)$
1	0		$A_1(0)$	$M_1(0)$	$L_1(0)$
1	1		$A_1(1)$	$M_1(1)$	$L_1(1)$
1
1	T-1	Y_1	$A_1(T-1)$	$M_1(T-1)$	$L_1(T-1)$
2	0		$A_2(0)$	$M_2(0)$	$L_2(0)$
2	1		$A_2(1)$	$M_2(1)$	$L_2(1)$
2
2	T-1	Y_2	$A_2(T-1)$	$M_2(T-1)$	$L_2(T-1)$
...
n	0		$A_n(0)$	$M_n(0)$	$L_n(0)$
n	1		$A_n(1)$	$M_n(1)$	$L_n(1)$
n
n	T-1	Y_n	$A_n(T-1)$	$M_n(T-1)$	$L_n(T-1)$

Table S.1-2. General form of person-time dataset when outcome is survival data ($S_i(t)$).

id	t	$S_i(t)$	$A_i(t)$	$M_i(t)$	$L_i(t)$
1	0	$S_1(0)$	$A_1(0)$	$M_1(0)$	$L_1(0)$
1	1	$S_1(1)$	$A_1(1)$	$M_1(1)$	$L_1(1)$
1
1	T-1	$S_1(T-1)$	$A_1(T-1)$	$M_1(T-1)$	$L_1(T-1)$
2	0	$S_2(0)$	$A_2(0)$	$M_2(0)$	$L_2(0)$
2	1	$S_2(1)$	$A_2(1)$	$M_2(1)$	$L_2(1)$
2
2	T-1	$S_2(T-1)$	$A_2(T-1)$	$M_2(T-1)$	$L_2(T-1)$
...
n	0	$S_n(0)$	$A_n(0)$	$M_n(0)$	$L_n(0)$
n	1	$S_n(1)$	$A_n(1)$	$M_n(1)$	$L_n(1)$
n
n	T-1	$S_n(T-1)$	$A_n(T-1)$	$M_n(T-1)$	$L_n(T-1)$

Table S.2. Summary of Outcome Types (otype).

cov1otype	Description	Modeled with
0	Linear change (or constant) *	Not modeled
1	Binary	Logistic regression model
2	Survival	Logistic regression model until failure
3	Continuous	Linear regression model

*The change for each time should be specified in *covlinc* parameter

Table S.3. Summary of Predictor Types (ptype).

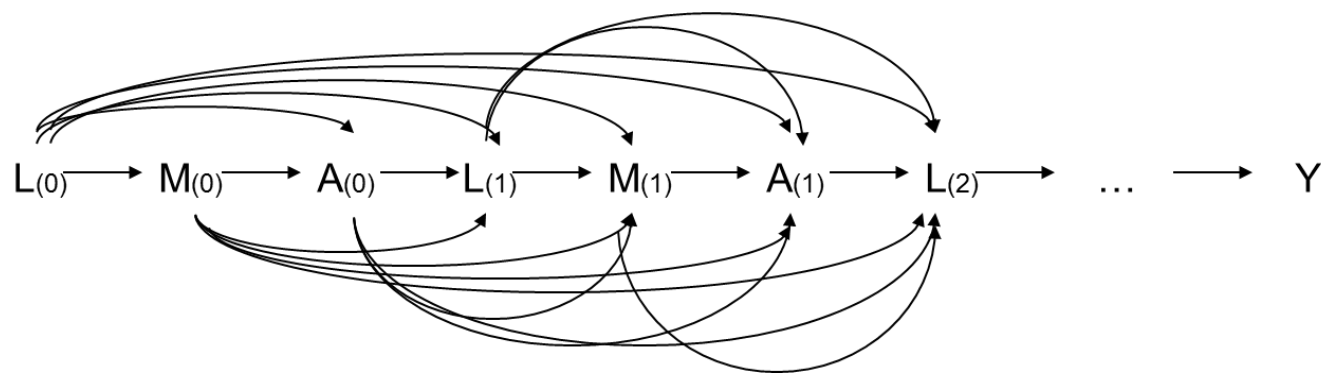
Cov1ptype	description
con-	Concurrent
lag1-	1 lagged period
lag2-	2 lagged period
lag3-	3 lagged period
-bin	Binary or linear
-qdc	Quadratic
-cub	Cubic

Table S.4. Output of the mgformula macro.

	Estimate	95% CI
$E[Y_{0G0}]$	135.691	134.93, 137.11
$E[Y_{1G0}]$	137.211	135.76, 138.80
$E[Y_{0G1}]$	135.336	134.57, 136.69
$E[Y_{1G1}]$	136.874	135.64, 138.37
rTE	1.18	-0.68, 2.69
rNDE	1.52	-0.25, 2.90
rNIE	-0.34	-0.52, -0.13

rTE: randomly interventional analogue of total effect; rDE: randomly interventional analogue of natural direct effect; rIE: randomly interventional analogue of natural indirect effect.

Figure S.1. Time-varying mediation with ordering of variables of $L(t)$, $M(t)$, and $A(t)$ for $t = 0$ to T .



Appendix 2.1. Identification of mediation parameter $\psi(a, a^*)$ to the survival mediational g-formula $Q(a, a^*)$

For simplicity of expression, we define survivor mediators $sM(t)^*$ and $sG(t)^*$ sequentially. First $sM(1)^*$ is defined as $M(1)_{a(1)^*}$ and $sG(1)^*$ as the random draw of $sM(1)^*$. For time t , the $sM(t)^*$ is defined as $M(1)_{a(1)^*, sG(1:t-1)^*, s(1:t-1)=1}$ and $sG(t)^*$ as the random draw of $sM(t)^*$. According to consistency assumption, when $A(1:t) = a(1:t)$, $M(1:t-1) = sM(t-1)$, and $S(t-1) = 1$, $sM(t)^* = M(t)$.

Lemma 1 *Given $A(1:t) = a(1:t)$, $M(1:t) = m(1:t)$, $S(1:t)^* = 1$, $sG(1:t)^* = m(1:t)$, (1) $sG(1:t)^* = G(1:t)^* = m(1:t)$ and $S(1:t) = S(1:t)^* = 1$, and (2) $A(t+1)$, $M(t+1)$, and $L(t+1)$ are all defined.*

step 1. $m(1) = sG(1)^* = G(1)^*$ (by definition)

step 2. $1 = S(1)^* = S(1)_{a(1), G(1)^*}$ (by definition)

$= S(1)_{a(1), m(1)}$ (by consistency)

$= S(1)$ (by consistency)

step 3. $m(2) = sG(2)^* = G(2)^*$ (since $S(1) = 1$)

step 4. $1 = S(2)^* = S(2)_{a(1:2), G(1:2)^*} = S(2)_{a(1:2), m(1:2)} = S(2)$ (by definition and consistency)

We can repeat step 3 and 4 iteratively for $t = 3, 4, \dots, t$, and demonstrate $sG(1:t)^* = G(1:t)^* = m(1:t)$ and $S(1:t) = S(1:t)^* = 1$. Because $S(t) = 1$, $A(t+1)$, $M(t+1)$, and $L(t+1)$ are all defined.

Proof for Identification of mediation parameter $\psi(a, a^*)$ to the survival mediational g-formula $Q(a, a^*)$

Under the following four sequential no unmeasured confounding assumptions for $t=1, 2, \dots, T$:

1. $S(T)_{a(1:T), m(1:t), s(1:t-1)=1, G(t+1:T)^*, S(t:T-1)^*} \perp A(t) | v, A(1:t-1) = a(1:t-1), M(1:t-1) = m(1:t-1), L(1:t-1), S(1:t-1)^* = 1, G(1:t)^* = m(1:t)$

$$2. S(T)_{a(1:T), m(1:t), s(1:t-1)=1, G(t+1:T)*, S(t:T-1)*} \perp M(t) | v, A(1:t) = a(1:t), M(1:t-1) = m(1:t-1), L(1:t-1), S(1:t-1)^* = 1, G(1:t)^* = m(1:t)$$

$$3. M(t)_{a(1:t)*, m(1:t-1), s(1:t-1)=1} \perp S(t-1) | A(1:t-1), M(1:t-1), L(1:t-1), S(1:t-2) = 1, v$$

$$4. M(t)_{a(1:t)*, m(1:t-1), s(1:t-1)=1} \perp A(t) | A(1:t-1), M(1:t-1), L(1:t-1), S(1:t-1) = 1, v$$

(In assumptions 1 and 2, the $A(t)$ and $M(t)$ are defined according to Lemma.)

$$\begin{aligned} \psi(a, a^*) &= \sum_v E[S(T)^* | v] \Pr(v) \\ E[S(T)^* | v] &\equiv E[S(T)_{a(1:T), G(1:T)*, S(1:T-1)*} | v] \\ &= \sum_{m(1)} E[S(T)_{a(1:T), m(1), G(2:T)*, S(1:T-1)*} | G(1)^* = m(1), v] \Pr(G(1)^* = m(1) | v) \quad (\text{add } G(1)) \\ &= \sum_{m(1)} E[S(T)_{a(1:T), m(1), G(2:T)*, S(1:T-1)*} | a(1), m(1), G(1)^* = m(1), v] \Pr(G(1)^* = m(1) | v) \\ &\quad (\text{add } a(1) \text{ and } m(1) \text{ since } A(1) \perp S(T)_{a(1:T), m(1), G(2:T)*, S(1:T-1)*} | G(1)^*, v) \\ &\quad \text{and } M(1) \perp S(T)_{a(1:T), m(1), G(2:T)*, S(1:T-1)*} | G(1)^*, A(1), v \text{ (assumptions 1 and 2 when } t=1)) \\ &= \sum_{m(1)l(1)} E[S(T)_{a(1:T), m(1), G(2:T)*, S(1:T-1)*} | a(1), m(1), l(1), G(1)^* = m(1), v] \\ &\quad \times \Pr(l(1) | a(1), m(1), v) \\ &\quad \times \Pr(sG(1)^* = m(1) | v) \\ &\quad (\text{add } L(1); \text{ by } G(1)^* = sG(1)^* \text{ and } l(1) \perp G(1)^* | A(1), M(1), v) \\ &= \sum_{m(1)l(1)} E[S(T)_{a(1:T), m(1), s(1)=1, G(2:T)*, S(2:T-1)*} | a(1), m(1), l(1), S(1)^* = 1, G(1)^* = m(1), v] \\ &\quad \times \Pr(S(1)^* = 1 | a(1), m(1), l(1), G(1)^* = m(1), v) \\ &\quad \times \Pr(l(1) | a(1), m(1), v) \\ &\quad \times \Pr(sG(1)^* = m(1) | v) \\ &\quad (\text{add } S(1)^*=1; \text{ remind } S(1)^* \equiv S(1)_{a(1), G(1)*} \text{ and } S(T)_{S(1)=0} = 0) \\ &= \sum_{m(1)l(1)} E[S(T)_{a(1:T), m(1), s(1)=1, G(2:T)*, S(2:T-1)*} | a(1), m(1), l(1), S(1)^* = 1, sG(1)^* = m(1), v] \\ &\quad \times E[S(1) | a(1), m(1), l(1), v] \end{aligned}$$

$$\begin{aligned}
& \times \Pr(l(1)|a(1), m(1), v) \\
& \times \Pr(sG(1)^* = m(1)|v) \\
& \text{(see proof 1-S1)} \\
& = \sum_{m(1:2)l(1)} E[S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)^*, S(2:T-1)^*} | a(1), m(1), l(1), S(1)^* = 1, sG(1)^* = \\
& m(1), G(2)^* = m(2), v] \\
& \times \Pr(G(2)^* = m(2)|a(1), m(1), l(1), G(1)^* = m(1), S(1)^* = 1, v) \\
& \times E[S(1)|a(1), m(1), l(1), v] \\
& \times \Pr(l(1)|a(1), m(1), v) \\
& \times \Pr(sG(1)^* = m(1)|v) \text{ (add } G(2)=m(2)) \\
& = \sum_{m(1:2)l(1)} E[S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)^*, S(2:T-1)^*} | a(1), m(1), l(1), S(1)^* = 1, sG(1:2)^* = \\
& m(1:2), v] \\
& \times \Pr(sG(2)^* = m(2)|a(1), m(1), l(1), sG(1)^* = m(1), S(1)^* = 1, v) \\
& \times E[S(1)|a(1), m(1), l(1), v] \\
& \times \Pr(l(1)|a(1), m(1), v) \\
& \times \Pr(sG(1)^* = m(1)|v) \text{ (by definition of } sG(2)^* \text{ since } S(1)^* = 1) \\
& = \sum_{m(1:2)l(1)} E[S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)^*, S(2:T-1)^*} | a(1), m(1), l(1), S(1)^* = 1, sG(1:2)^* = \\
& m(1:2), v] \\
& \times E[S(1)|a(1), m(1), l(1), v] \\
& \times \Pr(l(1)|a(1), m(1), v) \\
& \times \Pr(sG(1:2)^* = m(1:2)|v) \\
& \text{(see proof 1-G2)} \\
& = \sum_{m(1:2)l(1)} E[S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)^*, S(2:T-1)^*} | a(1:2), m(1:2), l(1), S(1)^* = 1, sG(1: \\
& 2)^* = m(1:2), v] \\
& \times E[S(1)|a(1), m(1), l(1), v] \\
& \times \Pr(l(1)|a(1), m(1), v) \\
& \times \Pr(sG(1:2)^* = m(1:2)|v) \\
& \text{(add A(2) and M(2))}
\end{aligned}$$

$$\begin{aligned}
& \text{since } S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)*, S(2:T-1)*} \perp A(2)|v, A(1), M(1), L(1), S(1)^* = 1, sG(1 : \\
& 2)^* = m(1 : 2) \text{ by assumption 1} \\
& \text{and } S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)*, S(2:T-1)*} \perp M(2)|v, A(1 : 2), M(1), L(1), S(1)^* = 1, sG(1 : \\
& 2)^* = m(1 : 2) \text{ by assumption 2)} \\
& = \sum_{m(1:2)l(1)} E[S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)*, S(2:T-1)*} | a(1 : 2), m(1 : 2), l(1 : 2), S(1)^* = \\
& 1, sG(1 : 2)^* = m(1 : 2), v] \\
& \quad \times \Pr(l(2)|a(1 : 2), m(1 : 2), l(1), S(1)^* = 1, sG(1 : 2)^* = m(1 : 2), v) \\
& \quad \times E(S(1)|a(1), m(1), l(1), v) \\
& \quad \times \Pr(l(1)|a(1), m(1), v) \\
& \quad \times \Pr(sG(1 : 2)^* = m(1 : 2)|v) \\
& \quad (\text{add L(2)}) \\
& = \sum_{m(1:2)l(1)} E[S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)*, S(2:T-1)*} | a(1 : 2), m(1 : 2), l(1 : 2), S(1)^* = \\
& 1, sG(1 : 2)^* = m(1 : 2), v] \\
& \quad \times \Pr(l(2)|a(1 : 2), m(1 : 2), l(1), S(1) = 1, v) \\
& \quad \times E(S(1)|a(1), m(1), l(1), v) \\
& \quad \times \Pr(l(1)|a(1), m(1), v) \\
& \quad \times \Pr(sG(1 : 2)^* = m(1 : 2)|v) \\
& \quad (\text{See proof 1-L2}) \\
& = \sum_{m(1:2)l(1)} E[S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)*, S(2:T-1)*} | a(1 : 2), m(1 : 2), l(1 : 2), S(1 : 2)^* = \\
& 1, sG(1 : 2)^* = m(1 : 2), v] \\
& \quad \times E[S(2)^* | a(1 : 2), m(1 : 2), l(1 : 2), S(1)^* = 1, sG(1 : 2)^* = m(1 : 2), v] \\
& \quad \times \Pr(l(2)|a(1 : 2), m(1 : 2), l(1), S(1) = 1, v) \\
& \quad \times E(S(1)|a(1), m(1), l(1), v) \\
& \quad \times \Pr(l(1)|a(1), m(1), v) \\
& \quad \times \Pr(sG(1 : 2)^* = m(1 : 2)|v) \\
& \quad (\text{add S(2)=1}) \\
& = \sum_{m(1:2)l(1)} E[S(T)_{a(1:T), m(1:2), s(1)=1, G(3:T)*, S(2:T-1)*} | a(1 : 2), m(1 : 2), l(1 : 2), S(1 : 2)^* =
\end{aligned}$$

$$\begin{aligned}
& 1, sG(1 : 2)^* = m(1 : 2), v] \\
& \times E[S(2)|a(1 : 2), m(1 : 2), l(1 : 2), S(1) = 1, v] \\
& \times \Pr(l(2)|a(1 : 2), m(1 : 2), l(1), S(1) = 1, v) \\
& \times E(S(1)|a(1), m(1), l(1), v) \\
& \times \Pr(l(1)|a(1), m(1), v) \\
& \times \Pr(sG(1 : 2)^* = m(1 : 2)|v) \\
& (\text{see proof 1-S2}) \\
& = \sum_{m(1:T)l(1:T)} E[S(T)_{a(1:T), m(1:T), s(T-1)=1} | a(1 : T), m(1 : T), l(1 : T), S(1 : T - 1)^* = \\
& 1, sG(1 : T)^* = m(1 : T), v] \\
& \times \prod_{t=1}^{T-1} E[S(t)|a(1 : t), m(1 : t), l(1 : t), S(t - 1) = 1, v] \\
& \times \prod_{t=1}^{T-1} \Pr(l(t)|a(1 : t), m(1 : t), l(1 : t - 1), S(1)^* = 1, v) \\
& \times \Pr(sG(1 : t)^* = m(1 : t)|v) \\
& (\text{repeat previous steps iteratively}) \\
& = \sum_{m(1:T)l(1:T)} E[S(T)|a(1 : T), m(1 : T), l(1 : T), S(1 : T - 1) = 1, v] \\
& \times \prod_{t=1}^{T-1} E[S(t)|a(1 : t), m(1 : t), l(1 : t), S(t - 1) = 1, v] \\
& \times \prod_{t=1}^{T-1} \Pr(l(t)|a(1 : t), m(1 : t), l(1 : t - 1), S(1)^* = 1, v) \\
& \times \Pr(sG(1 : T)^* = m(1 : T)|v) \\
& (\text{see proof 1-ST}) \\
& = \sum_{m(1:T)l(1:T)} E[S(T)|a(1 : T), m(1 : T), l(1 : T), S(1 : T - 1) = 1, v] \\
& \times \prod_{t=1}^{T-1} E[S(t)|a(1 : t), m(1 : t), l(1 : t), S(t - 1) = 1, v] \\
& \times \prod_{t=1}^{T-1} \Pr(l(t)|a(1 : t), m(1 : t), l(1 : t - 1), S(1)^* = 1, v) \\
& \times \sum_{l^\dagger(1:T-1)} \prod_{t=1}^T \Pr(m(t)|a(1 : t)^*, m(1 : t - 1), l^\dagger(1 : t - 1), S(t - 1) = 1, v) \\
& \times \Pr(l^\dagger(t - 1)|a^*(1 : t - 1), m(1 : t - 1), l^\dagger(1 : t - 2), S(t - 2) = 1, v) \\
& (\text{see proof 1-GT}) \\
& \psi(a, a^*) = \sum_v E[S(T)^*|v] \Pr(v) \\
& = \sum_{v, m(1:T), l(1:T)} E[S(T)|a(1 : T), m(1 : T), l(1 : T), S(1 : T - 1) = 1, v] \\
& \times \prod_{t=1}^{T-1} E[S(t)|a(1 : t), m(1 : t), l(1 : t), S(t - 1) = 1, v]
\end{aligned}$$

$$\begin{aligned}
& \times \prod_{t=1}^{T-1} \Pr(l(t)|a(1:t), m(1:t), l(1:t-1), S(1)^* = 1, v) \\
& \times \sum_{l^\dagger(1:T-1)} \prod_{t=1}^T \Pr(m(t)|a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), S(t-1) = 1, v) \\
& \times \Pr(l^\dagger(t-1)|a^*(1:t-1), m(1:t-1), l^\dagger(1:t-2), S(t-2) = 1, v) \\
& \times \Pr(v)
\end{aligned}$$

which is the survival mediational g-formula, $Q(a, a^*)$.

Proof 1-S1:

$$\begin{aligned}
& \Pr(S(1)^* = 1|a(1), m(1), l(1), G(1)^* = m(1), v) \\
& = E(S(1)_{a(1), G(1)^*} | a(1), m(1), l(1), G(1)^* = m(1), v) \text{ (by definition)} \\
& = E(S(1)_{a(1), m(1), l(1)} | a(1), m(1), l(1), G(1)^* = m(1), v) \text{ (consistency)} \\
& = E(S(1)_{a(1), m(1), l(1)} | a(1), m(1), l(1), v) \text{ (since } S(1)_{a(1), m(1), l(1)} \perp G(1)^* | A(1), M(1), L(1), v) \\
&) \\
& = E(S(1) | a(1), m(1), l(1), v) \text{ (consistency)}
\end{aligned}$$

Proof 1-G2:

$$\begin{aligned}
& \Pr(sG(2)^* = m(2)|a(1), m(1), l(1), sG(1)^* = m(1), S(1)^* = 1, v) \times \Pr(sG(1)^* = m(1)|v) \\
& = \frac{\Pr(sG(2)^* = m(2), sG(1)^* = m(1)|a(1), m(1), l(1), S(1)^* = 1, v)}{\Pr(sG(1)^* = m(1)|a(1), m(1), l(1), S(1)^* = 1, v)} \times \Pr(sG(1)^* = m(1)|v) \\
& = \frac{\Pr(sG(2)^* = m(2), sG(1)^* = m(1)|v)}{\Pr(sG(1)^* = m(1)|v)} \times \Pr(sG(1)^* = m(1)|v) \\
& \text{(since } sG(1)^* \perp A(1), M(1), L(1), S(1)|v \text{ and } (sG(1)^*, sG(2)^*) \perp A(1), M(1), L(1), S(1)|v) \\
& = \Pr(sG(2)^* = m(2), sG(1)^* = m(1)|v) = \Pr(sG(1:2)^* = m(1:2)|v) \text{ (by definition)}
\end{aligned}$$

Proof 1-L2

$$\begin{aligned}
& \Pr(l(2)|a(1:2), m(1:2), l(1), S(1)^* = 1, sG(1:2)^* = m(1:2), v) \\
& = \Pr(l(2)|a(1:2), m(1:2), l(1), S(1) = 1, sG(1:2)^* = m(1:2), v) \quad \text{(Lemma)} \\
& = \Pr(l(2)|a(1:2), m(1:2), l(1), S(1) = 1, v) \text{ (since } sG(1:2)^* \perp l(2) | A(1:2), M(1:2), L(1), S(1) = 1, v)
\end{aligned}$$

Proof 1-S2

$$\begin{aligned}
& E[S(2)^* | a(1:2), m(1:2), l(1:2), S(1)^* = 1, sG(1:2)^* = m(1:2), v] \\
& = E[S(2) | a(1:2), m(1:2), l(1:2), S(1) = 1, sG(1:2)^* = m(1:2), v] \quad \text{(Lemma)} \\
& = E[S(2) | a(1:2), m(1:2), l(1:2), S(1) = 1, v] \text{ (since } sG(1:2)^* \perp S(2) | A(1:2), M(1:2), L(1), S(1) = 1, v)
\end{aligned}$$

2), $L(1 : 2), S(1) = 1, v$)

Proof 1-ST

$$\begin{aligned}
& E[S(T)_{a(1:T), m(1:T), s(T-1)=1} | a(1 : T), m(1 : T), l(1 : T), S(1 : T - 1)^* = 1, sG(1 : T)^* = \\
& m(1 : T), v] \\
& = E[S(T)_{a(1:T), m(1:T), s(T-1)=1} | a(1 : T), m(1 : T), l(1 : T), S(1 : T - 1) = 1, sG(1 : T)^* = \\
& m(1 : T), v] \quad (\text{Lemma}) \\
& = E[S(T) | a(1 : T), m(1 : T), l(1 : T), S(1 : T - 1) = 1, v] \quad (\text{since } sG(1 : T)^* \perp \\
& S(T) | A(1 : T), M(1 : T), L(1 : T), S(1 : T - 1) = 1, v)
\end{aligned}$$

Proof 1-GT

$$\begin{aligned}
& \Pr(sG(1 : T)^* = m(1 : T) | v) \\
& = \prod_{t=1}^T \Pr(sG(1 : t)^* = m(t) | v) / \Pr(sG(1 : t - 1)^* = m(1 : t - 1), v) \\
& = \prod_{t=1}^T \Pr(sG(1 : t)^* = m(t) | a(1 : t - 1)^*, m(1 : t - 1), l^\dagger(1 : t - 2), S(1 : t - 2) = \\
& 1, v) / \Pr(sG(1 : t - 1)^* = m(1 : t - 1) | a(1 : t - 1)^*, m(1 : t - 1), l^\dagger(1 : t - 2), S(1 : t - 2) = \\
& 1, v) (\text{since } sG(1 : t)^* \perp (A(1 : t - 1), M(1 : t - 1), L(1 : t - 2), S(1 : t - 2)) | v) \\
& = \prod_{t=1}^T \Pr(sG(t)^* = m(t) | a(1 : t - 1)^*, m(1 : t - 1), l^\dagger(1 : t - 2), S(1 : t - 2) = 1, sG(1 : \\
& t - 1)^* = m(1 : t - 1), v) \\
& = \prod_{t=1}^T \Pr(sM(t)^* = m(t) | a(1 : t - 1)^*, m(1 : t - 1), l^\dagger(1 : t - 2), S(1 : t - 2) = 1, sG(1 : \\
& t - 1)^* = m(1 : t - 1), v) \quad (\text{definition of } sG(t)^*) \\
& = \prod_{t=1}^T \Pr(M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} = m(t) | a(1 : t - 1)^*, m(1 : t - 1), l^\dagger(1 : t - 2), S(1 : \\
& t - 2) = 1, sG(1 : t - 1)^* = m(1 : t - 1), v) \quad (\text{definition of } sM(t)^* \text{ and consistency}) \\
& = \prod_{t=1}^T \sum_{l^\dagger(t)} \Pr((M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} = m(t) | a(1 : t - 1)^*, m(1 : t - 1), l^\dagger(1 : \\
& t - 1), S(1 : t - 2) = 1, sG(1 : t - 1)^* = m(1 : t - 1), v) \Pr(l^\dagger(t - 1) | a(1 : t - 1)^*, m(1 : \\
& t - 1), l^\dagger(1 : t - 2), S(1 : t - 2) = 1, v) \quad (\text{add } L(t) \text{ and } sG(1 : t - 1)^* \perp L(t - 1) | A(1 : \\
& t - 1), M(1 : t - 1) = sG(1 : t - 1)^* = m(1 : t - 1), L(1 : t - 2), S(1 : t - 2) = 1, v) \\
& = \prod_{t=1}^T \sum_{l^\dagger(t)} \Pr((M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} = m(t) | a(1 : t - 1)^*, m(1 : t - 1), l^\dagger(1 : \\
& t - 1), S(1 : t - 1) = 1, sG(1 : t - 1)^* = m(1 : t - 1), v) \Pr(l^\dagger(t - 1) | a(1 : t - 1)^*, m(1 : t - 1), l^\dagger(1 : \\
& t - 2), S(1 : t - 2) = 1, v) \quad (\text{add } S(t-1) \text{ and } M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} \perp S(t - 1) | A(1 :
\end{aligned}$$

$$\begin{aligned}
& t-1), M(1:t-1), L(1:t-1), S(1:t-2) = 1, v \text{ by assumption 3)} \\
& = \prod_{t=1}^T \sum_{l^\dagger(t)} \Pr(M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} = m(t) | a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), S(1:t-1) = 1, sG(1:t-1)^* = m(1:t-1), v) \Pr(l^\dagger(t-1) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), S(1:t-2) = 1, v) \text{ (add } A(t) \text{ by } M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} \perp A(t) | A(1:t-1), M(1:t-1) = sG(1:t-1)^* = m(1:t-1), L(1:t-1), S(1:t-1) = 1, v \text{ by assumption 4)} \\
& = \prod_{t=1}^T \sum_{l^\dagger(t)} \Pr(M(t) = m(t) | a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), S(1:t-1) = 1, sG(1:t-1)^* = m(1:t-1), v) \Pr(l^\dagger(t-1) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), S(1:t-2) = 1, v) \\
& \text{(by consistency)} \\
& = \prod_{t=1}^T \sum_{l^\dagger(t)} \Pr(m(t) | a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), S(1:t-1) = 1, v) \Pr(l^\dagger(t-1) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), S(1:t-2) = 1, v) \text{ (} sG(1:t-1)^* \perp M(t) | A(1:t), M(1:t-1), L(1:t-1), S(1:t-1) = 1)
\end{aligned}$$

Appendix 2.2. Identification for traditional definition of mediation parameter

$\Phi(a, a^*)$ under the sequential monotonicity assumption

In order to identify the traditional definition of mediation parameter $\Phi(a, a^*)$, we should make (1) monotonicity assumption, (2) four no unmeasured confounding assumptions, and (3) crossworld independence assumption and no Under the monotonicity assumption as below.

(1) monotonicity assumption

$$S(t)_{a^*, m(1:t)} \geq S(t)_{a, m(1:t)} \text{ where } t \in \{1, 2, \dots, T-1\}$$

(2) the five sequential no unmeasured confounding assumptions for $t=1, 2, \dots, T$:

$$1. S(t)_{a, m(1:t), S(t-1)=1} \perp A | v$$

$$2. S(t)_{a, m(1:t), S(t-1)=1} \perp M(k) | v, A, M(1:k-1), S(k-1)=1 \text{ where } k=1 \dots t$$

$$3. M(t)_{a^* m(1:t-1) s(t-1)=1} \perp S(t-1) | A, M(1:t-1)$$

$$4. M(t)_{a^* m(1:t-1) s(t-1)=1} \perp A | v$$

which are similar to the four assumptions in Appendix 1;

(3) a crossworld independence assumption

$$S(t)_{a, m(1:t), S(t-1)=1} \perp \{M(k)_{a^* m(k-1) s(k-1)=1}, k=1 \dots t\} | v$$

which holds only under settings without time-varying confounders.

If the monotonicity assumption holds, the $\Phi(a, a^*)$ can be defined. If the above assumptions all hold, we can identify the $\Phi(a, a^*)$ as the following expression:

$$\begin{aligned} & \sum_{m(1)} E[S(T)_{a, m(1:T), S(T-1)=1} | v, M(k)_{a^* m(k-1) s(k-1)=1} = m(k), k = \{1, 2, \dots, T\}, S(T-1)_{a, m(1:T-1)} = 1] \\ & \prod_{t=1}^T E[S(t) | v, a, m(1:t), S(t-1)=1] \prod_{t=1}^T \Pr(M(t) = m(t) | A = a^*, m(1:t-1), S(t-1) = 1) \end{aligned}$$

$$\Phi(a, a^*) \equiv E[S(T)_{aM(1:T)a^*} | v] \equiv E[S(T)_{aM(1:T)a^*S(1:T-1)a^*} S(1:T-1)_{a, M(1:T-1)a^*} | v]$$

$$\begin{aligned} & = \sum_{m(1)} E[S(T)_{am(1)M(2:T)a^*m(1)S(1:T-1)a^*m(1)} S(1:T-1)_{a, m(1), M(2:T-1)a^*} \\ & | v, M(1)_{a^*} = m(1)] \end{aligned}$$

$$\Pr(M(1)_{a^*} = m(1) | v) \text{ (add } M(1)a^*)$$

$$= \sum_{m(1)} E[S(T)_{am(1)M(2:T)a^*m(1)S(1:T-1)a^*m(1)} S(1:T-1)_{a, m(1), M(2:T-1)a^*}$$

$$\begin{aligned}
& |v, M(1)_{a^*} = m(1)] \\
& \Pr(M(1) = m(1)|v, A = a^*) \quad (M(1)_a \perp A|v \text{ and consistency}) \\
& = \sum_{m(1)} E[S(T)_{am(1), S(1)=1, M(2:T)_{a^*m(1)S(1:T-1)_{a^*m(1)}}} S(2:T-1)_{a, m(1), M(2:T-1)_{a^*}} \\
& |v, M(1)_{a^*} = m(1), S(1)_{a, m(1)} = 1] \\
& E[S(1)_{a, m(1)}|v, M(1)_{a^*} = m(1)] \Pr(M(1) = m(1)|v, A = a^*) \quad (\text{add } S(1)) \\
& = \sum_{m(1)} E[S(T)_{am(1), S(1)=1, M(2:T)_{a^*m(1)S(1:T-1)_{a^*m(1)}}} S(2:T-1)_{a, m(1), M(2:T-1)_{a^*}} \\
& |v, M(1)_{a^*} = m(1), S(1)_{a, m(1)} = 1] \\
& E[S(1)|v, a, m(1)] \Pr(M(1) = m(1)|v, A = a^*) \\
& (S(1)_{a, m(1)} \perp M(1)_{a^*}, S(1)_{a, m(1)} \perp A|v, S(1)_{a, m(1)} \perp M|v, A, \text{and consistency,}) \\
& = \sum_{m(1)} E[S(T)_{am(1), S(1)=1, M(2)_{a^*m(1)S(1)}_{a^*m(1)}} M(3:T)_{a^*m(1)S(1:T-1)_{a^*m(1)}} S(2:T-1)_{a, m(1), M(2:T-1)_{a^*}} \\
& |v, M(1)_{a^*} = m(1), S(1)_{a, m(1)} = 1] \\
& E[S(1)|v, a, m(1)] \Pr(M(1) = m(1)|v, A = a^*) \quad (\text{by definition of } M(2:T)) \\
& = \sum_{m(1)} E[S(T)_{am(1), S(1)=1, M(2)_{a^*m(1)s(1)=1}} M(3:T)_{a^*m(1), s(1)=1, S(2:T-1)_{a^*m(1)}} S(2:T-1)_{a, m(1), M(2:T-1)_{a^*m(1)s(1)=1}} \\
& |v, M(1)_{a^*} = m(1), S(1)_{a, m(1)} = 1] \\
& E[S(1)|v, a, m(1)] \Pr(M(1) = m(1)|v, A = a^*) \quad (\text{by monotonicity assumption } S(1)_{a^*m(1)} \geq \\
& S(1)_{a, m(1)}) \\
& = \sum_{m(1)} E[S(T)_{a, m(1:2), S(1)=1, M(3:T)_{a^*m(1:2), s(1)=1, S(2:T-1)_{a^*m(1:2)}}} S(2:T-1)_{a, m(1:2), M(3:T-1)_{a^*m(1:2)s(1)=1}} \\
& |v, M(1)_{a^*} = m(1), S(1)_{a, m(1)} = 1, M(2)_{a^*m(1)s(1)=1} = m(2)] \\
& \Pr(M(2)_{a^*m(1)s(1)=1} = m(2)|M(1)_{a^*} = m(1)) E[S(1)|v, a, m(1)] \Pr(M(1) = m(1)|v, A = \\
& a^*) \quad (\text{add } M(2)) \\
& = \sum_{m(1)} E[S(T)_{a, m(1:2), S(1)=1, M(3:T)_{a^*m(1:2), s(1)=1, S(2:T-1)_{a^*m(1:2)}}} S(2:T-1)_{a, m(1:2), M(3:T-1)_{a^*m(1:2)s(1)=1}} \\
& |v, M(1)_{a^*} = m(1), S(1)_{a, m(1)} = 1, M(2)_{a^*m(1)s(1)=1} = m(2)] \\
& \Pr(M(2) = m(2)|A = a^*, M(1) = m(1), S(1) = 1) E[S(1)|v, a, m(1)] \Pr(M(1) = m(1)|v, A = \\
& a^*) \\
& (M(2)_{a^*m(1)s(1)=1} \perp A, \text{consistency, and } M(2)_{a^*m(1)s(1)=1} \perp S(1)|A, M(1), \text{and consistency} \\
& \text{again}) \\
& = \sum_{m(1)} E[S(T)_{a, m(1:2), S(2)=1, M(3:T)_{a^*m(1:2), s(1)=1, S(2:T-1)_{a^*m(1:2)}}} S(3:T-1)_{a, m(1:2), M(3:T-1)_{a^*m(1:2)s(1)=1}} |
\end{aligned}$$

$$\begin{aligned}
& v, M(1)_{a^*} = m(1), S(1)_{a,m(1)} = 1, M(2)_{a^*m(1)s(1)=1} = m(2), S(2)_{a,m(1:2)} = 1] \\
& E[S(2)_{a,m(1:2)}|v, M(1)_{a^*} = m(1), S(1)_{a,m(1)} = 1, M(2)_{a^*m(1)s(1)=1} = m(2)] \\
& \Pr(M(2) = m(2)|A = a^*, M(1) = m(1), S(1) = 1))E[S(1)|v, a, m(1)] \\
& \Pr(M(1) = m(1)|v, A = a^*) \quad (\text{add } S(2)) \\
& = \sum_{m(1)} E[S(T)_{a,m(1:2),S(2)=1,M(3:T)_{a^*m(1:2),s(1:2)=1,S(3:T-1)_{a^*m(1:2)}}^{S(3:T-1)_{a,m(1:2),M(3:T-1)_{a^*m(1:2)s(1:2)=1}}}] \\
& v, M(1)_{a^*} = m(1), S(1)_{a,m(1)} = 1, M(2)_{a^*m(1)s(1)=1} = m(2), S(2)_{a,m(1:2)} = 1] \\
& E[S(2)_{a,m(1:2)}|v, M(1)_{a^*} = m(1), S(1)_{a,m(1)} = 1, M(2)_{a^*m(1)s(1)=1} = m(2)] \\
& \Pr(M(2) = m(2)|A = a^*, M(1) = m(1), S(1) = 1)) \\
& E[S(1)|v, a, m(1)] \Pr(M(1) = m(1)|v, A = a^*) \\
& (\text{by monotonicity assumption } S(2)_{a^*,m(1:2)} \geq S(2)_{a,m(1:2)}) \\
& = \sum_{m(1)} E[S(T)_{a,m(1:2),S(2)=1,M(3:T)_{a^*m(1:2),s(1:2)=1,S(3:T-1)_{a^*m(1:2)}}^{S(3:T-1)_{a,m(1:2),M(3:T-1)_{a^*m(1:2)s(1:2)=1}}}] \\
& v, M(1)_{a^*} = m(1), S(1)_{a,m(1)} = 1, M(2)_{a^*m(1)s(1)=1} = m(2), S(2)_{a,m(1:2)} = 1] \\
& E[S(2)|v, A = a, M(1:2) = m(1:2), S(1) = 1] \\
& \Pr(M(2) = m(2)|A = a^*, M(1) = m(1), S(1) = 1)) \\
& E[S(1)|v, a, m(1)] \Pr(M(1) = m(1)|v, A = a^*) \\
& (S(2)_{a,m(1:2)} \perp (A, M(1)_{a^*}, M(2)_{a^*m(1)s(1)=1})|S(1)_{a,m(1)} = 1, v; S(2)_{a,m(1:2)} \perp M(1)|A, v; \text{consistency;} \\
& S(2)_{a,m(1:2)} \perp M(2)|A, M(1), S(1) = 1, v; \text{ and consistency}) \\
& = \sum_{m(1)} E[S(T)_{a,m(1:2),S(2)=1,M(3:T)_{a^*m(1:2),s(1:2)=1,S(3:T-1)_{a^*m(1:2)}}^{S(3:T-1)_{a,m(1:2),M(3:T-1)_{a^*m(1:2)s(1:2)=1}}}] \\
& v, M(1)_{a^*} = m(1), S(1)_{a,m(1)} = 1, M(2)_{a^*m(1)s(1)=1} = m(2), S(2)_{a,m(1:2)} = 1] \\
& \prod_{t=1}^2 E[S(t)|v, a, m(1:t), S(t-1) = 1] \\
& \prod_{t=1}^2 \Pr(M(t) = m(t)|A = a^*, m(1:t-1), S(t-1) = 1)) \quad (\text{change the notation}) \\
& \dots(\text{repeat the above procedure}) \\
& = \sum_{m(1)} E[S(T)_{a,m(1:T),S(T-1)=1}|v, M(k)_{a^*m(k-1)s(k-1)=1} = m(k), k \in \{1, 2, \dots, T\}, S(T-1)_{a,m(1:T-1)} = 1] \\
& \prod_{t=1}^T E[S(t)|v, a, m(1:t), S(t-1) = 1] \prod_{t=1}^T \Pr(M(t) = m(t)|A = a^*, m(1:t-1), S(t-1) = 1))
\end{aligned}$$

$$\begin{aligned}
&= \sum_{m(1)} E[S(T)|v, a, m(1:T), S(T-1) = 1] \prod_{t=1}^T E[S(t)|v, a, m(1:t), S(t-1) = 1] \\
&\prod_{t=1}^T \Pr(M(t) = m(t)|A = a^*, m(1:t-1), S(t-1) = 1) \\
&(S(T)_{a,m(1:T),S(T-1)=1} \perp A|v; S(T)_{a,m(1:T),S(T-1)=1} \perp M(t)|v, A, M(1:t-1), S(t-1) = 1;
\end{aligned}$$

and consistency)

$$\begin{aligned}
&= \sum_{m(1)} E[S(T)_{a,m(1:T),S(T-1)=1}|v, M(k)_{a^*m(k-1)s(k-1)=1} = m(k), k = \{1, 2, \dots, T\}, S(T-1)_{a,m(1:T-1)} = 1] \\
&\prod_{t=1}^T E[S(t)|v, a, m(1:t), S(t-1) = 1] \\
&\prod_{t=1}^T \Pr(M(t) = m(t)|A = a^*, m(1:t-1), S(t-1) = 1) \\
&(S(T)_{a,m(1:T),S(T-1)=1} \perp \{M(k)_{a^*m(k-1)s(k-1)=1}, k = 1 \dots T\}|v)
\end{aligned}$$

Appendix 2.3 Stochastic process version for $\psi(a, a^*)$ to the survival mediational g-formula $Q(a, a^*)$

In this section, the stochastic counting process notation, $N(t)$, is used to replace the survival variable $S(t)$. We also treat time-varying mediators and confounders as continuous variables. According to the definition of counting process, $N(t) = 1 - S(t)$. The identical definition for $S(t)^*$ also applies to the definition of $N(t)^*$.

The mediation parameter, $\psi(a, a^*)$, is defined as $\Pr[N(T)^* = 0|v]$. Under the following four sequential no unmeasured confounding assumptions for $t=1, 2, \dots, T$:

$$1. N(T)_{a(1:T), m(1:t), N(t-1)=1, G(t+1:T)^*, N(t:T-1)^*} \perp A(t)|v, A(1:t-1) = a(1:t-1), M(1:t-1) = m(1:t-1), L(1:t-1), N(1:t-1)^* = 0, G(1:t)^* = m(1:t)$$

$$2. N(T)_{a(1:T), m(1:t), N(t-1)=1, G(t+1:T)^*, N(t:T-1)^*} \perp M(t)|v, A(1:t) = a(1:t), M(1:t-1) = m(1:t-1), L(1:t-1), N(1:t-1)^* = 1, G(1:t)^* = m(1:t)$$

$$3. M(t)_{a(1:t)^*, m(1:t-1), N(t-1)=0} \perp N(t-1)|A(1:t-1), M(1:t-1), L(1:t-1), N(t-2) = 0, v$$

$$4. M(t)_{a(1:t)^*, m(1:t-1), N(t-1)=0} \perp A(t)|A(1:t-1), M(1:t-1), L(1:t-1), N(t-1) = 0, v$$

(In assumptions 1 and 2, the $A(t)$ and $M(t)$ are defined according to Lemma.),

$\psi(a, a^*)$ can be identified as the survival mediational g-formula, $Q(a, a^*)$,

where $Q(a, a^*)$

$$\begin{aligned} &= \int_{m(1:T)l(1:T)} E[1 - N(T)|a(1:T), m(1:T), l(1:T), N(1:T-1) = 0, v] \\ &\times \prod_{t=1}^{T-1} E[1 - N(t)|a(1:t), m(1:t), l(1:t), N(t-1) = 0, v] \\ &\times \prod_{t=1}^{T-1} \int_{L(t)} f_{L(t)}(l(t)|a(1:t), m(1:t), l(1:t-1), N(t-1) = 0, v) dl(1:T) \\ &\times \int_{l^\dagger(1:T-1)} \prod_{t=1}^T f_{M(t)}(m(t)|a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), N(t-1) = 0, v) \\ &\times f_{L^\dagger(t-1)}(l^\dagger(t-1)|a^*(1:t-1), m(1:t-1), l^\dagger(1:t-2), N(t-2) = 0, v) dm(1:T) dl^\dagger(1:T-1) \end{aligned}$$

The proof is as follows.

Lemma 2 Given $A(1:t) = a(1:t)$, $M(1:t) = m(1:t)$, $N(1:t)^* = 0$, $sG(1:t)^* = m(1:t)$, (1) $sG(1:t)^* = G(1:t)^* = m(1:t)$ and $N(1:t) = N(1:t)^* = 0$, and (2) $A(t+1)$, $M(t+1)$, and $L(t+1)$ are all defined.

step 1. $m(1) = sG(1)^* = G(1)^*$ (by definition)

step 2. $0 = N(1)^* = N(1)_{a(1),G(1)^*}$ (by definition)

$= N(1)_{a(1),m(1)}$ (by consistency)

$= N(1)$ (by consistency)

step 3. $m(2) = sG(2)^* = G(2)^*$ (since $N(1) = 0$)

step 4. $0 = N(2)^* = N(2)_{a(1:2),G(1:2)^*} = N(2)_{a(1:2),m(1:2)} = N(2)$ (by definition and consistency)

We can repeat step 3 and 4 iteratively for $t = 3, 4, \dots, t$, and demonstrate $sG(1:t)^* = G(1:t)^* = m(1:t)$ and $N(1:t) = N(1:t)^* = 0$. Because $N(t) = 0$, $A(t+1)$, $M(t+1)$, and $L(t+1)$ are all defined.

Proof for Identification of mediation parameter $\psi(a, a^*)$ to the survival mediational g-formula $Q(a, a^*)$

Under the following four sequential no unmeasured confounding assumptions for $t=1, 2, \dots, T$:

1. $N(T)_{a(1:T),m(1:t),N(t-1)=1,G(t+1:T)^*,N(t:T-1)^*} \perp A(t)|v, A(1:t-1) = a(1:t-1), M(1:t-1) = m(1:t-1), L(1:t-1), N(1:t-1)^* = 0, G(1:t)^* = m(1:t)$

2. $N(T)_{a(1:T),m(1:t),N(t-1)=1,G(t+1:T)^*,N(t:T-1)^*} \perp M(t)|v, A(1:t) = a(1:t), M(1:t-1) = m(1:t-1), L(1:t-1), N(1:t-1)^* = 1, G(1:t)^* = m(1:t)$

3. $M(t)_{a(1:t)^*,m(1:t-1),N(t-1)=0} \perp N(t-1)|A(1:t-1), M(1:t-1), L(1:t-1), N(t-2) = 0, v$

4. $M(t)_{a(1:t)^*,m(1:t-1),N(t-1)=0} \perp A(t)|A(1:t-1), M(1:t-1), L(1:t-1), N(t-1) = 0, v$

(In assumptions 1 and 2, the $A(t)$ and $M(t)$ are defined according to Lemma.)

$\psi(a, a^*) \equiv E[S(T)^*|v] \equiv \Pr[N(T)^* = 0|v] = E[1 - N(T)^*|v] \equiv E[1 - N(T)_{a(1:T),G(1:T)^*,N(1:T-1)^*}|v]$
 $= \int_{m(1)} E[1 - N(T)_{a(1:T),m(1),G(2:T)^*,N(1:T-1)^*}|G(1)^* = m(1), v] f_{G(1)^*}(m(1)|v) d m(1)$ (add $G(1)$)

$= \int_{m(1)} E[1 - N(T)_{a(1:T),m(1),G(2:T)^*,N(1:T-1)^*}|a(1), m(1), G(1)^* = m(1), v] f_{G(1)^*}(m(1)|v) d m(1)$ (add $a(1)$ and $m(1)$ since $A(1) \perp N(T)_{a(1:T),m(1),G(2:T)^*,N(1:T-1)^*}|G(1)^*, v$ and $M(1) \perp N(T)_{a(1:T),m(1),G(2:T)^*,N(1:T-1)^*}|G(1)^*, A(1), v$ (assumptions 1 and 2 when $t=1$))

$$\begin{aligned}
&= \int_{m(1)l(1)} E[1 - N(T)_{a(1:T),m(1),G(2:T)*,N(1:T-1)*} | a(1), m(1), l(1), G(1)^* = m(1), v] \\
&\quad f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1)*}(m(1) | v) d l(1) dm(1) \\
&\quad (\text{add } L(1); \text{ by } G(1)^* = sG(1)^* \text{ and } l(1) \perp G(1)^* | A(1), M(1), v) \\
&= \int_{m(1)l(1)} E[1 - N(T)_{a(1:T),m(1),N(1)=0,G(2:T)*,N(2:T-1)*} | a(1), m(1), l(1), N(1)^* = 0, G(1)^* = \\
&\quad m(1), v] \\
&\quad \times \Pr(N(1)^* = 0 | a(1), m(1), l(1), G(1)^* = m(1), v) f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1)*}(m(1) | v) d \\
&\quad l(1) dm(1) \\
&\quad (\text{add } N(1)^* = 0; \text{ remind } N(1)^* \equiv N(1)_{a(1),G(1)*} \text{ and } N(T)_{N(1)=1} = 1, E[1 - N(T)_{N(1)=1}] = \\
&\quad 0) \\
&= \int_{m(1)l(1)} E[1 - N(T)_{a(1:T),m(1),N(1)=0,G(2:T)*,N(2:T-1)*} | a(1), m(1), l(1), N(1)^* = 0, sG(1)^* = \\
&\quad m(1), v] \\
&\quad \times E[1 - N(1) | a(1), m(1), l(1), v] f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1)*}(m(1) | v) d l(1) dm(1) \\
&\quad (\text{see proof 1-N1}) \\
&= \int_{m(1:2)l(1)} E[1 - N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} | a(1), m(1), l(1), N(1)^* = 0, sG(1)^* = \\
&\quad m(1), G(2)^* = m(2), v] \\
&\quad \times f_{G(2)*}(m(2) | a(1), m(1), l(1), G(1)^* = m(1), N(1)^* = 0, v) \\
&\quad \times E[1 - N(1) | a(1), m(1), l(1), v] f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1)*}(m(1) | v) d l(1) dm(1 : 2) \quad (\text{add} \\
&\quad G(2)=m(2)) \\
&= \int_{m(1:2)l(1)} E[1 - N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} | a(1), m(1), l(1), N(1)^* = 0, sG(1 : \\
&\quad 2)^* = m(1 : 2), v] \\
&\quad \times f_{sG(2)*}(m(2) | a(1), m(1), l(1), sG(1)^* = m(1), N(1)^* = 0, v) \\
&\quad \times E[1 - N(1) | a(1), m(1), l(1), v] f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1)*}(m(1) | v) d l(1) dm(1 : 2) \quad (\text{by} \\
&\quad \text{definition of } sG(2)^* \text{ since } N(1)^* = 0) \\
&= \int_{m(1:2)l(1)} E[1 - N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} | a(1), m(1), l(1), N(1)^* = 0, sG(1 : \\
&\quad 2)^* = m(1 : 2), v] \\
&\quad \times E[1 - N(1) | a(1), m(1), l(1), v] \\
&\quad \times f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1:2)*}(m(1 : 2) | v) d l(1) dm(1 : 2)
\end{aligned}$$

(see proof 1-G2)

$$= \int_{m(1:2)l(1)} E[1 - N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} | a(1 : 2), m(1 : 2), l(1), N(1)^* = 0, sG(1 : 2)^* = m(1 : 2), v]$$

$$\times E[1 - N(1) | a(1), m(1), l(1), v]$$

$$\times f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1:2)*}(m(1 : 2) | v) d l(1) dm(1 : 2)$$

(add A(2) and M(2))

$$\text{since } N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} \perp A(2) | v, A(1), M(1), L(1), N(1)^* = 0, sG(1 : 2)^* = m(1 : 2) \text{ by assumption 1}$$

$$\text{and } N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} \perp M(2) | v, A(1 : 2), M(1), L(1), N(1)^* = 0, sG(1 : 2)^* = m(1 : 2) \text{ by assumption 2)}$$

$$= \int_{m(1:2)l(1)} E[1 - N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} | a(1 : 2), m(1 : 2), l(1 : 2), N(1)^* = 0, sG(1 : 2)^* = m(1 : 2), v]$$

$$\times f_{L(2)}(l(2) | a(1 : 2), m(1 : 2), l(1), N(1)^* = 0, sG(1 : 2)^* = m(1 : 2), v)$$

$$\times E[1 - N(1) | a(1), m(1), l(1), v]$$

$$\times f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1:2)*}(m(1 : 2) | v) d l(1 : 2) dm(1 : 2)$$

(add L(2))

$$= \int_{m(1:2)l(1)} E[1 - N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} | a(1 : 2), m(1 : 2), l(1 : 2), N(1)^* = 0, sG(1 : 2)^* = m(1 : 2), v]$$

$$\times f_{L(2)}(l(2) | a(1 : 2), m(1 : 2), l(1), N(1) = 0, v)$$

$$\times E[1 - N(1) | a(1), m(1), l(1), v]$$

$$\times f_{L(1)}(l(1) | a(1), m(1), v) f_{sG(1:2)*}(m(1 : 2) | v) d l(1 : 2) dm(1 : 2)$$

(See proof 1-L2)

$$= \int_{m(1:2)l(1)} E[1 - N(T)_{a(1:T),m(1:2),N(1)=0,G(3:T)*,N(2:T-1)*} | a(1 : 2), m(1 : 2), l(1 : 2), N(1 : 2)^* = 0, sG(1 : 2)^* = m(1 : 2), v]$$

$$\times E[1 - N(2)^* | a(1 : 2), m(1 : 2), l(1 : 2), N(1)^* = 0, sG(1 : 2)^* = m(1 : 2), v]$$

$$\times f_{L(2)}(l(2) | a(1 : 2), m(1 : 2), l(1), N(1) = 0, v)$$

$$\times E[1 - N(1) | a(1), m(1), l(1), v]$$

$$\begin{aligned}
& \times f_{L(1)}(l(1)|a(1), m(1), v) f_{sG(1:2)^*}(m(1:2)|v) d l(1:2) dm(1:2) \\
& (\text{add } N(2)^*=0) \\
& = \int_{m(1:2)l(1)} E[1 - N(T)_{a(1:T), m(1:2), N(1)=0, G(3:T)^*, N(2:T-1)^*} | a(1:2), m(1:2), l(1:2), N(1:2)^* = 0, sG(1:2)^* = m(1:2), v] \\
& \times E[1 - N(2)|a(1:2), m(1:2), l(1:2), N(1)^* = 0, sG(1:2)^* = m(1:2), v] \\
& \times f_{L(2)}(l(2)|a(1:2), m(1:2), l(1), N(1) = 0, v) \\
& \times E[1 - N(1)|a(1), m(1), l(1), v] \\
& \times f_{L(1)}(l(1)|a(1), m(1), v) f_{sG(1:2)^*}(m(1:2)|v) d l(1:2) dm(1:2) \\
& (\text{see proof 1-N2}) \\
& = \int_{m(1:T)l(1:T)} E[1 - N(T)_{a(1:T), m(1:T), N(T-1)=0} | a(1:T), m(1:T), l(1:T), N(1:T-1)^* = 0, sG(1:T)^* = m(1:T), v] \\
& \times \prod_{t=1}^{T-1} E[1 - N(t)|a(1:t), m(1:t), l(1:t), N(t-1) = 0, v] \\
& \times \prod_{t=1}^{T-1} f_{L(t)}(l(t)|a(1:t), m(1:t), l(1:t-1), N(t-1) = 0, v) \\
& \times f_{sG(1:t)^*}(m(1:t)|v) d l(1:T) dm(1:T) \\
& (\text{repeat previous steps iteratively}) \\
& = \int_{m(1:T)l(1:T)} E[1 - N(T)|a(1:T), m(1:T), l(1:T), N(1:T-1) = 0, v] \\
& \times \prod_{t=1}^{T-1} E[1 - N(t)|a(1:t), m(1:t), l(1:t), N(t-1) = 0, v] \\
& \times \prod_{t=1}^{T-1} f_{L(t)}(l(t)|a(1:t), m(1:t), l(1:t-1), N(t-1) = 0, v) \\
& \times f_{sG(1:t)^*}(m(1:t)|v) d l(1:T) dm(1:T) \\
& (\text{see proof 1-NT}) \\
& = \int_{m(1:T)l(1:T)} E[1 - N(T)|a(1:T), m(1:T), l(1:T), N(1:T-1) = 0, v] \\
& \times \prod_{t=1}^{T-1} E[1 - N(t)|a(1:t), m(1:t), l(1:t), N(t-1) = 0, v] \\
& \times \prod_{t=1}^{T-1} f_{L(t)}(l(t)|a(1:t), m(1:t), l(1:t-1), N(t-1) = 0, v) d l(1:T) \\
& \times \int_{l^\dagger(1:T-1)} \prod_{t=1}^T f_{M(t)}(m(t)|a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), N(t-1) = 0, v) \\
& \times f_{L^\dagger(t-1)}(l^\dagger(t-1)|a^*(1:t-1), m(1:t-1), l^\dagger(1:t-2), N(t-2) = 0, v) dm(1:T) d \\
& l^\dagger(1:T-1) \\
& (\text{see proof 1-GT})
\end{aligned}$$

which is the survival mediational g-formula, $Q(a, a^*)$.

Proof 1-N1:

$$\begin{aligned}
& \Pr(N(1)^* = 0 | a(1), m(1), l(1), G(1)^* = m(1), v) \\
&= E(1 - N(1)_{a(1), G(1)^*} | a(1), m(1), l(1), G(1)^* = m(1), v) \text{ (by definition)} \\
&= E(1 - N(1)_{a(1), m(1), l(1)} | a(1), m(1), l(1), G(1)^* = m(1), v) \text{ (consistency)} \\
&= E(1 - N(1)_{a(1), m(1), l(1)} | a(1), m(1), l(1), v) \text{ (since } N(1)_{a(1), m(1), l(1)} \perp G(1)^* | A(1), M(1), L(1), v) \\
& \quad) \\
&= E(1 - N(1) | a(1), m(1), l(1), v) \text{ (consistency)}
\end{aligned}$$

Proof 1-G2:

$$\begin{aligned}
& f_{sG(2)^*}(m(2) | a(1), m(1), l(1), sG(1)^* = m(1), N(1)^* = 0, v) \times f_{sG(1)^*}(m(1) | v) \\
&= \frac{f_{sG(2)^*}(m(2), sG(1)^* = m(1) | a(1), m(1), l(1), N(1)^* = 0, v)}{f_{sG(1)^*}(m(1) | a(1), m(1), l(1), N(1)^* = 0, v)} \times f_{sG(1)^*}(m(1) | v) \\
&= \frac{f_{sG(2)^*}(m(2), sG(1)^* = m(1) | v)}{f_{sG(1)^*}(m(1) | v)} \times f_{sG(1)^*}(m(1) | v) \\
& \text{(since } sG(1)^* \perp A(1), M(1), L(1), N(1) | v \text{ and } (sG(1)^*, sG(2)^*) \perp A(1), M(1), L(1), N(1) | v) \\
&= f(sG(2)^* = m(2), sG(1)^* = m(1) | v) = f_{sG(1:2)^*}(m(1:2) | v) \text{ (by definition)}
\end{aligned}$$

Proof 1-L2

$$\begin{aligned}
& f_{L(2)}(l(2) | a(1:2), m(1:2), l(1), N(1)^* = 0, sG(1:2)^* = m(1:2), v) \\
&= f_{L(2)}(l(2) | a(1:2), m(1:2), l(1), N(1) = 0, sG(1:2)^* = m(1:2), v) \quad \text{(Lemma)} \\
&= f_{L(2)}(l(2) | a(1:2), m(1:2), l(1), N(1) = 0, v) \text{ (since } sG(1:2)^* \perp l(2) | A(1:2), M(1:2), L(1), N(1) = 0, v)
\end{aligned}$$

Proof 1-N2

$$\begin{aligned}
& E[N(2)^* | a(1:2), m(1:2), l(1:2), N(1)^* = 0, sG(1:2)^* = m(1:2), v] \\
&= E[N(2) | a(1:2), m(1:2), l(1:2), N(1) = 0, sG(1:2)^* = m(1:2), v] \quad \text{(Lemma)} \\
&= E[N(2) | a(1:2), m(1:2), l(1:2), N(1) = 0, v] \text{ (since } sG(1:2)^* \perp N(2) | A(1:2), M(1:2), L(1:2), N(1) = 0, v)
\end{aligned}$$

Proof 1-NT

$$\begin{aligned}
& E[N(T)_{a(1:T), m(1:T), s(T-1)=1} | a(1:T), m(1:T), l(1:T), N(1:T-1)^* = 0, sG(1:T)^* = m(1:T), v]
\end{aligned}$$

$$\begin{aligned}
&= E[N(T)_{a(1:T), m(1:T), s(T-1)=1} | a(1:T), m(1:T), l(1:T), N(1:T-1) = 0, sG(1:T)^* = m(1:T), v] \quad (\text{Lemma}) \\
&= E[N(T) | a(1:T), m(1:T), l(1:T), N(1:T-1) = 0, v] \quad (\text{since } sG(1:T)^* \perp N(T) | A(1:T), M(1:T), L(1:T), N(1:T-1) = 0, v)
\end{aligned}$$

Proof 1-GT

$$\begin{aligned}
&f_{sG(1:T)^*}(m(1:T) | v) \\
&= \prod_{t=1}^T f_{sG(t)^*}(m(t) | v) / f_{sG(1:t-1)^*}(m(1:t-1) | v) \\
&= \prod_{t=1}^T f_{sG(t)^*}(m(t) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, v) / f_{sG(1:t-1)^*}(m(1:t-1) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, v) (\text{since } sG(1:t)^* \perp (A(1:t-1), M(1:t-1), L(1:t-2), N(1:t-2)) | v) \\
&= \prod_{t=1}^T f_{sG(t)^*}(m(t) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, sG(1:t-1)^* = m(1:t-1), v) \\
&= \prod_{t=1}^T f_{sM(t)^*}(m(t) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, sG(1:t-1)^* = m(1:t-1), v) \quad (\text{definition of } sG(t)^*) \\
&= \prod_{t=1}^T f_{M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1}}(m(t) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, sG(1:t-1)^* = m(1:t-1), v) \quad (\text{definition of } sM(t)^* \text{ and consistency}) \\
&= \prod_{t=1}^T \int_{l^\dagger(t)} f_{M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1}}(m(t) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-1), N(1:t-2) = 0, sG(1:t-1)^* = m(1:t-1), v) f_{L(t-1)}(l^\dagger(t-1) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, v) dl^\dagger(t-1) \quad (\text{add } L(t) \text{ and } sG(1:t-1)^* \perp L(t-1) | A(1:t-1), M(1:t-1) = sG(1:t-1)^* = m(1:t-1), L(1:t-2), N(1:t-2) = 0, v) \\
&= \prod_{t=1}^T \int_{l^\dagger(t)} f_{M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1}}(m(t) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-1), N(1:t-1) = 0, sG(1:t-1)^* = m(1:t-1), v) f_{L(t-1)}(l^\dagger(t-1) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, v) dl^\dagger(t-1) \quad (\text{add } N(t-1) \text{ and } M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} \perp N(t-1) | A(1:t-1), M(1:t-1), L(1:t-1), N(1:t-2) = 0, v \text{ by assumption 3}) \\
&= \prod_{t=1}^T \int_{l^\dagger(t)} f_{M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1}}(m(t) | a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), N(1:t-1) = 0, sG(1:t-1)^* = m(1:t-1), v) f_{L(t-1)}(l^\dagger(t-1) | a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, v) dl^\dagger(t-1) \quad (\text{add } A(t) \text{ by } M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} \perp A(t) | A(1:t-1), M(1:t-1), L(1:t-1), N(1:t-2) = 0, v)
\end{aligned}$$

$$\begin{aligned}
& t-1) = sG(1:t-1)^* = m(1:t-1), L(1:t-1), N(1:t-1) = 0, v \text{ by assumption 4)} \\
& = \prod_{t=1}^T \int_{l^\dagger(t)} f_{M(t)}(m(t)|a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), N(1:t-1) = 0, sG(1:t-1)^* = \\
& m(1:t-1), v) f_{L(t-1)}(l^\dagger(t-1)|a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, v) dl^\dagger(t-1) \\
& \text{(by consistency)} \\
& = \prod_{t=1}^T \int_{l^\dagger(t)} f_{M(t)}(m(t)|a(1:t)^*, m(1:t-1), l^\dagger(1:t-1), N(1:t-1) = 0, v) f_{L(t-1)}(l^\dagger(t-1)|a(1:t-1)^*, m(1:t-1), l^\dagger(1:t-2), N(1:t-2) = 0, v) dl^\dagger(t-1) (sG(1:t-1)^* \perp M(t)|A(1:t), M(1:t-1), L(1:t-1), N(1:t-1) = 0)
\end{aligned}$$

Appendix 2.4. Assumption 3 with unmeasured survival confounders

In section 3, we mentioned that the four assumptions required for effect identification allow the unmeasured confounding between survival statuses at different time points. To justify this argument, we adapt the non-parametric structural equation model (NPSEM) to generate one setting with unmeasured survival confounders in which the assumption 3 is still held.

We describe the causal mechanisms for all variables in Figure 1 by NPSEM as follows:

$$V = g_V(\varepsilon_V)$$

$$A(1) = g_{A_1}(V, \varepsilon_{A_1})$$

$$M(1) = g_{M_1}(V, A(1), \varepsilon_{M_1})$$

$$L(1) = g_{L_1}(V, A(1), M(1), \varepsilon_{L_1})$$

$$S(1) = g_{S_1}(V, A(1), M(1), L(1), \varepsilon_{S_1})$$

and for $t=2, \dots, T$

$A(t) = g_{A_t}(V, A(1:t-1), M(1:t-1), L(1:t-1), \varepsilon_{A_t})$ when $S(t-1) = 1$; undefined when $S(t-1) = 0$

$M(t) = g_{M_t}(V, A(1:t), M(1:t-1), L(1:t-1), \varepsilon_{M_t})$ when $S(t-1) = 1$; undefined when $S(t-1) = 0$

$L(t) = g_{L_t}(V, A(1:t), M(1:t), L(1:t-1), \varepsilon_{L_t})$ when $S(t-1) = 1$; undefined when $S(t-1) = 0$

$$S(t) = g_{S_t}(V, A(1:t), M(1:t), L(1:t), S(t-1), \varepsilon_{S_t})$$

For convenience, we rewrite the assumption 3 by assigning values for the conditioned random variables, as follows:

$$M(t)_{a^*(1:t), m(1:t-1), s(1:t-1)=1} \perp S(t-1) | V = v, A(1:t-1) = a_c(1:t-1), M(1:t-1) = m_c(1:t-1), L(1:t-1) = l_c(1:t-1), S(1:t-2) = 1$$

Under NPSEM and that condition,

$$M(t)_{a^*(1:t), m(1:t-1), s(1:t-1)=1} = g_{M_t}(V, a^*(1:t), m(1:t-1), l_c(1:t-1), \varepsilon_{M_t})$$

$$S(t-1) = g_{S_{t-1}}(v, a_c(1:t-1), m_c(1:t-1), l_c(1:t-1), S(t-2) = 1, \varepsilon_{S_{t-1}})$$

It is obvious that ε_{M_t} and $\varepsilon_{S_{t-1}}$ are the only random parts of $M(t)_{a*(1:t),m(1:t-1),s(1:t-1)=1}$ and $S(t-1)$, respectively. Therefore, the independence assumption, $\varepsilon_{M_t} \perp \varepsilon_{S_{t-1}} | S(t-2) = 1$, is sufficient to ensure that the assumption 3 is not violated. A further assumption of no unmeasured survival confounders (i.e. $(\varepsilon_{S_1}, \dots, \varepsilon_{S_t}, \dots, \varepsilon_{S_T})$ are mutually independent) is not required.

Appendix 3.1. Identification of randomly interventional analogue of mediation parameter $r\Phi(a, a', a'', a''')$

Proof A

$$\begin{aligned}
r\Phi(a, a', a'', a''') &\equiv E[Y(a, G_1(a'), G_2(a'', G_1(a''')))] \\
&= \sum_{m_1} E[Y(a, m_1, G_2(a'', G_1(a''')))|G_1(a') = m_1] \Pr(G_1(a') = m_1) \text{ (add } G_1(a')) \\
&= \sum_{m_1} E[Y(a, m_1, G_2(a'', G_1(a''')))] \Pr(G_1(a') = m_1) \\
&\quad (G_1(a') \text{ depends to all variables, including } G_1(a''')) \\
&= \sum_{m_1, m_2} E[Y(a, m_1, m_2)|G_2(a'', G_1(a''')) = m_2] \Pr(G_2(a'', G_1(a''')) = m_2) \Pr(G_1(a') = m_1) \\
&\quad (\text{add } G_2(a'', G_1(a'''))) \\
&= \sum_{m_1, m_2} E[Y(a, m_1, m_2)] \Pr(G_2(a'', G_1(a''')) = m_2) \Pr(G_1(a') = m_1) \\
&\quad (G_2(a'', G_1(a''')) \text{ depends to all variables}) \\
&= \sum_{m_1, m_2} E[Y(a, m_1, m_2)] \Pr(M_2(a'', G_1(a''')) = m_2) \Pr(M_1(a') = m_1) \\
&\quad (G \text{ and } M \text{ have the same distribution}) \\
&= \sum_{m_1, m_2, m'_1} E[Y(a, m_1, m_2)] \Pr(M_2(a'', m'_1) = m_2 | G_1(a''') = m'_1) \\
&\quad \Pr(G_1(a''') = m'_1) \Pr(M_1(a') = m_1) \\
&\quad (\text{add } G_1(a''')) \\
&= \sum_{m_1, m_2, m'_1} E[Y(a, m_1, m_2)] \Pr(M_2(a'', m'_1) = m_2) \Pr(G_1(a''') = m'_1) \Pr(M_1(a') = m_1) \\
&\quad (G_1(a') \text{ depends to all variables}) \\
&= \sum_{m_1, m_2, m'_1} E[Y(a, m_1, m_2)] \Pr(M_2(a'', m'_1) = m_2) \Pr(M_1(a''') = m'_1) \Pr(M_1(a') = m_1) \\
&\quad (G \text{ and } M \text{ have the same distribution})
\end{aligned}$$

Proof B

All components in the last equation of proof a can be identified as the expression of g-formula. In order to do that, we need to make four no unmeasured confounding assumptions: (1) no unmeasured exposure-outcome confounding ($Y(a, m_1, m_2) \perp A|V$), (2) no unmeasured mediator-outcome confounding ($Y(a, m_1, m_2) \perp M_1|V, A, L_1$ and $Y(a, m_1, m_2) \perp M_2|V, A, L_1, M_1, L_2$), (3) no unmeasured exposure-mediator confounding ($(M_1(a), M_2(a, m_1)) \perp$

$A|V$), and (4) no unmeasured mediator-mediator confounding ($M_2(a, m_1) \perp M_1|V, A, L_1$).

Under the four assumptions, all components can be identified in forms of g-formula as follows.

$$E[Y(a, m_1, m_2)|v] = \sum_{l_1, l_2} E[Y|v, a, l_1, m_1, l_2, m_2] \Pr(l_1|v, a) \Pr(l_2|v, a, l_1, m_1)$$

$$\Pr(M_2(a, m_1) = m_2|v) = \sum_{l_1} \Pr(m_2|v, a, m_1, l_1) \Pr(l_1|v, a)$$

$$\Pr(M_1(a) = m_1|v) = \Pr(m_1|v, a)$$

Based on the above three expressions,

$$\begin{aligned} & \sum_{m_1, m_2, m'_1} E[Y(a, m_1, m_2)] \Pr(M_1(a') = m_1) \Pr(M_2(a'', m'_1) = m_2) \Pr(M_1(a''') = m'_1) \\ &= \sum_{v, m_1, m_2, m'_1} E[Y(a, m_1, m_2)|v] \Pr(M_1(a') = m_1|v) \Pr(M_2(a'', m'_1) = m_2|v) \Pr(M_1(a''') = m'_1|v) \Pr(v) \\ &= \sum_{v, m_1, m_2, l_1, l_2} E[Y|v, a, l_1, m_1, l_2, m_2] \Pr(l_1|v, a) \Pr(l_2|v, a, l_1, m_1) \Pr(m_1|v, a') \\ & \quad \sum_{l'_1, m'_1} \Pr(m_2|v, a'', m'_1, l'_1) \Pr(l'_1|v, a'') \Pr(m'_1|v, a''') \Pr(v) \end{aligned}$$

Proof C

$$\begin{aligned} r\Phi(a, a', a'', a') &\equiv E[Y(a, G_1(a'), G_2(a'', G_1(a')))] \\ &= \sum_{m_1} E[Y(a, m_1, G_2(a'', G_1(a')))|G_1(a') = m_1] \Pr(G_1(a') = m_1) \text{ (add } G_1(a')) \\ &= \sum_{m_1} E[Y(a, m_1, G_2(a'', m_1))] \Pr(G_1(a') = m_1) \\ &\text{(} G_1(a') \text{ depends to all variables)} \\ &= \sum_{m_1, m_2} E[Y(a, m_1, m_2)|G_2(a'', m_1) = m_2] \Pr(G_2(a'', m_1) = m_2) \Pr(G_1(a') = m_1) \\ &\text{(add } G_2(a'', m_1)) \\ &= \sum_{m_1, m_2} E[Y(a, m_1, m_2)] \Pr(G_2(a'', m_1) = m_2) \Pr(G_1(a') = m_1) \\ &\text{(} G_2(a'', m_1) \text{ depends to all variables)} \\ &= \sum_{m_1, m_2} E[Y(a, m_1, m_2)] \Pr(M_2(a'', m_1) = m_2) \Pr(M_1(a') = m_1) \\ &\text{(G and M have the same distribution)} \\ &= \sum_{v, m_1, m_2, l_1, l_2} E[Y|v, a, l_1, m_1, l_2, m_2] \Pr(l_1|v, a) \Pr(l_2|v, a, l_1, m_1) \Pr(m_1|v, a') \\ & \quad \sum_{l'_1} \Pr(m_2|v, a'', m_1, l'_1) \Pr(l'_1|v, a'') \Pr(v) \end{aligned}$$

(similar to proof B)

When time-varying confounders V , L_1 and L_2 are empty, then

$$= \sum_{m_1, m_2} E[Y|v, a, m_1, m_2] \Pr(m_1|v, a') \Pr(m_2|v, a'', m_1)$$

Appendix 3.2

Consider settings with one exposure (A), k mediators (M_1, M_2, \dots, M_k), and k sets of mediator-outcome confounder (L_1, L_2, \dots, L_k). For simplicity we assume that the exposure A was randomly assigned and there is no baseline confounder. Since the number of rPSEs is $2k$ and there is no close form for all rPSEs, here we use $k = 3$ as example to illustrate the methods.

When $k = 3$, the number of rPSEs is $23 = 8$, which are listed as follows:

$$\begin{aligned}
rPSE_{A \rightarrow Y} &= r\Phi(a_1, a_0, a_0, a_0, a_0, a_0, a_0, a_0) - r\Phi(a_0, a_0, a_0, a_0, a_0, a_0, a_0, a_0) \\
rPSE_{A \rightarrow M_1 \rightarrow Y} &= r\Phi(a_1, a_1, a_0, a_0, a_0, a_0, a_0, a_0) - r\Phi(a_1, a_0, a_0, a_0, a_0, a_0, a_0, a_0) \\
rPSE_{A \rightarrow M_2 \rightarrow Y} &= r\Phi(a_1, a_1, a_1, a_0, a_0, a_0, a_0, a_0) - r\Phi(a_1, a_1, a_0, a_0, a_0, a_0, a_0, a_0) \\
rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y} &= r\Phi(a_1, a_1, a_1, a_1, a_0, a_0, a_0, a_0) - r\Phi(a_1, a_1, a_1, a_0, a_0, a_0, a_0, a_0) \\
rPSE_{A \rightarrow M_3 \rightarrow Y} &= r\Phi(a_1, a_1, a_1, a_1, a_1, a_0, a_0, a_0) - r\Phi(a_1, a_1, a_1, a_1, a_0, a_0, a_0, a_0) \\
rPSE_{A \rightarrow M_1 \rightarrow M_3 \rightarrow Y} &= r\Phi(a_1, a_1, a_1, a_1, a_1, a_1, a_0, a_0) - r\Phi(a_1, a_1, a_1, a_1, a_1, a_0, a_0, a_0) \\
rPSE_{A \rightarrow M_2 \rightarrow M_3 \rightarrow Y} &= r\Phi(a_1, a_1, a_1, a_1, a_1, a_1, a_1, a_0) - r\Phi(a_1, a_1, a_1, a_1, a_1, a_1, a_0, a_0) \\
rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow Y} &= r\Phi(a_1, a_1, a_1, a_1, a_1, a_1, a_1, a_1) - r\Phi(a_1, a_1, a_1, a_1, a_1, a_1, a_1, a_0)
\end{aligned}$$

$r\Phi(a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)}, a^{(5)}, a^{(6)}, a^{(7)}, a^{(8)})$ is defined as

$$E[Y(a^{(1)}, G_1(a^{(2)}), G_2(a^{(3)}, G_1(a^{(4)}))G_3(a^{(5)}, G_1(a^{(6)}), G_2(a^{(7)}, G_1(a^{(8)})))).$$

Using the similar procedure in Appendix 3.1, $r\Phi(a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)}, a^{(5)}, a^{(6)}, a^{(7)}, a^{(8)}) =$

$$\sum_{m_3} \left[\begin{aligned} &\sum_{m_2} \left\{ \sum_{m_1} E[Y(a^{(1)}, m_1, m_2, m_3)] \times \Pr(M_1(a^{(2)}) = m_1) \right\} \\ &\times \left\{ \sum_{m'_1} \Pr[M_2(a^{(3)}, m'_1) = m_2] \times \Pr(M_1(a^{(4)}) = m'_1) \right\} \\ &\sum_{m''_1, m'''_1, m'_2} \Pr[M_3(a^{(5)}, m''_1, m'_2) = m_3] \times \Pr(M_1(a^{(6)}) = m''_1) \\ &\times \Pr[M_2(a^{(7)}, m'''_1) = m'_2] \times \Pr(M_1(a^{(8)}) = m'''_1) \end{aligned} \right].$$

Under no unmeasured confounding assumptions, each component of above expression can be identified by the g-formula as follows.

$$\begin{aligned}
E[Y(a, m_1, m_2, m_3)] &= \sum_{l_1, l_2, l_3} E[Y|a, l_1, m_1, l_2, m_2, l_3, m_3] \times \Pr(l_1|a) \times \Pr(l_2|a, l_1, m_1) \times \\
&\Pr(l_3|a, l_1, m_1, l_2, m_2) \\
\Pr(M_1(a) = m_1) &= \Pr(m_1|a)
\end{aligned}$$

$$\Pr[M_2(a, m_1) = m_2] = \sum_{l_1} \Pr(m_2|a, l_1, m_1) \Pr(l_1|a)$$

$$\Pr[M_3(a, m_1, m_2) = m_3] = \sum_{l_1, l_2} \Pr(m_3|a, l_1, m_1, l_2, m_2) \Pr(l_1|a) \Pr(l_2|a, l_1, m_1)$$

We can extend the above formula to any k by the similar procedure.

Appendix 3.3. A regression-based approach

In Appendix 1, we showed that under four no unmeasured confounding assumptions, $r\Phi(a, a', a'', a''')$ can be identified as $Q(a, a', a'', a''')$. In this section, we provide a regression-based approach for all rPSEs. Consider the mediator-outcome confounders are not affected by exposure, i.e. L_1 and L_2 are empty, the randomly interventional analogue of mediation parameter $Q(a, a', a'', a''')$ reduces to

$$\sum_{m_1, m_2, m'_1} E[Y|c, a, m_1, m_2] \Pr(m_1|c, a') \Pr(m_2|c, a'', m'_1) \Pr(m'_1|c, a''')$$

Then consider linear model for Y , M_1 , and M_2 with interaction term as follows (we omit high-order interaction term $\theta_7 am_1 m_2$ for simplicity).

$$E[Y|A = a, M_1 = m_1, M_2 = m_2, C = c] = \theta_0 + \theta_1 a + \theta_2 m_1 + \theta_3 m_2 + \theta_4 am_1 + \theta_5 am_2 + \theta_6 m_1 m_2 + \theta_c c$$

$$E[M_2|A = a, M_1 = m_1, C = c] = \beta_0 + \beta_1 a + \beta_2 m_1 + \beta_3 am_1 + \beta_c c$$

$$E[M_1|A = a, C = c] = \gamma_0 + \gamma_1 a + \gamma_c c.$$

Therefore,

$$\begin{aligned} Q(a, a', a'', a''') &= \sum_{m_1, m_2, m'_1} E[Y|a, m_1, m_2] \Pr(m_1|a') \Pr(m_2|a'', m'_1) \Pr(m'_1|a''') \\ &= \sum_{m_1, m_2, m'_1} (\theta_0 + \theta_1 a + \theta_2 m_1 + \theta_3 m_2 + \theta_4 am_1 + \theta_5 am_2 + \theta_6 m_1 m_2 + \theta_c c) \Pr(m_1|a') \\ &\quad \Pr(m_2|a'', m'_1) \Pr(m'_1|a''') \\ &= \theta_0 + \theta_c c + \theta_1 a + (\theta_2 + \theta_4 a) \sum_{m_1} m_1 \Pr(m_1|a') \\ &\quad + (\theta_3 + \theta_5 a) \sum_{m_2, m'_1} m_2 \Pr(m_2|a'', m'_1) \Pr(m'_1|a''') \\ &\quad + \theta_6 \sum_{m_1} m_1 \Pr(m_1|a') \sum_{m_2, m'_1} m_2 \Pr(m_2|a'', m'_1) \Pr(m'_1|a''') \\ &= \theta_0 + \theta_c c + \theta_c c + \theta_1 a + (\theta_2 + \theta_4 a) E[M_1|A = a'] \\ &\quad + ((\theta_3 + \theta_5 a) \sum_{m'_1} E[M_2|A = a'', M_1 = m'_1] \Pr(m'_1|a''') \\ &\quad + \theta_6 E[M_1|A = a'] \sum_{m_2, m'_1} E[M_2|A = a'', M_1 = m'_1] \Pr(m'_1|a''')) \\ &= \theta_0 + \theta_c c + \theta_1 a + (\theta_2 + \theta_4 a)(\gamma_0 + \gamma_1 a' + \gamma_c c) \\ &\quad + ((\theta_3 + \theta_5 a) \sum_{m'_1} E[M_2|A = a'', M_1 = m'_1] \Pr(m'_1|a''') \\ &\quad + \theta_6(\gamma_0 + \gamma_1 a' + \gamma_c c) \sum_{m_2, m'_1} E[M_2|A = a'', M_1 = m'_1] \Pr(m'_1|a''')) \end{aligned}$$

$$\begin{aligned}
&= \theta_0 + \theta_c c + \theta_1 a + (\theta_2 + \theta_4 a)(\gamma_0 + \gamma_1 a' + \gamma_c c) \\
&+ [(\theta_3 + \theta_5 a + \theta_6(\gamma_0 + \gamma_1 a' + \gamma_c c)) \sum_{m'_1} E[M_2|A = a'', M_1 = m'_1] \Pr(m'_1|a''')] \\
&= \theta_0 + \theta_c c + \theta_1 a + (\theta_2 + \theta_4 a)(\gamma_0 + \gamma_1 a' + \gamma_c c) \\
&+ [(\theta_3 + \theta_5 a + \theta_6(\gamma_0 + \gamma_1 a' + \gamma_c c)) \sum_{m'_1} (\beta_0 + \beta_1 a'' + \beta_2 m'_1 + \beta_3 a'' m'_1 + \beta_c c) \Pr(m'_1|a''')] \\
&= \theta_0 + \theta_c c + \theta_1 a + (\theta_2 + \theta_4 a)(\gamma_0 + \gamma_1 a' + \gamma_c c) \\
&+ [(\theta_3 + \theta_5 a + \theta_6(\gamma_0 + \gamma_1 a' + \gamma_c c))[(\beta_0 + \beta_1 a'' + \beta_c c) + (\beta_2 + \beta_3 a'')E[M_1|A = a''']] \\
&= \theta_0 + \theta_c c + \theta_1 a + (\theta_2 + \theta_4 a)(\gamma_0 + \gamma_1 a' + \gamma_c c) \\
&+ [(\theta_3 + \theta_5 a + \theta_6(\gamma_0 + \gamma_1 a' + \gamma_c c))[(\beta_0 + \beta_1 a'' + \beta_c c) + (\beta_2 + \beta_3 a'')(\gamma_0 + \gamma_1 a''' + \gamma_c c)]] \\
&= [\theta_0 + \theta_c c + \theta_2(\gamma_0 + \gamma_c c) + (\theta_3 + \theta_6 \gamma_0 + \theta_6 \gamma_c c)(\beta_0 + \beta_2 \gamma_0 + \beta_2 \gamma_c c + \beta_c c)] \\
&+ [\theta_1 + \theta_5(\beta_0 + \beta_2 \gamma_0 + \beta_2 \gamma_c c + \beta_c c) + \theta_4(\gamma_0 + \gamma_c c)]a \\
&+ [\theta_2 \gamma_1 + \theta_6 \gamma_1(\beta_0 + \beta_2 \gamma_0 + \beta_2 \gamma_c c + \beta_c c) + \theta_6 \gamma_1(\beta_0 + \beta_2 \gamma_0 + \beta_2 \gamma_c c + \beta_c c)]a' \\
&+ (\theta_3 + \theta_6 \gamma_0 + \theta_6 \gamma_c c)(\beta_1 + \beta_3 \gamma_0 + \beta_3 \gamma_c c)a'' + (\theta_3 + \theta_6 \gamma_0 + \theta_6 \gamma_c c)\beta_2 \gamma_1 a''' \\
&+ \theta_4 \gamma_1 a a' + \theta_5(\beta_1 + \beta_3 \gamma_0 + \beta_3 \gamma_c c)a a'' + \theta_5 \beta_2 \gamma_1 a a''' \\
&+ \theta_6 \gamma_1(\beta_1 + \beta_3 \gamma_0 + \beta_3 \gamma_c c)a' a'' + \theta_6 \beta_2 \gamma_1^2 a' a''' + (\theta_3 + \theta_6 \gamma_0 + \theta_6 \gamma_c c)\beta_3 \gamma_1 a'' a''' \\
&+ \theta_5 \beta_3 \gamma_1 a a'' a''' + \theta_6 \beta_3 \gamma_1^2 a' a'' a'''
\end{aligned}$$

and

$$\begin{aligned}
rPSE_{A \rightarrow Y} &= Q(a_1, a_0, a_0, a_0) - Q(a_0, a_0, a_0, a_0) \\
&= \left\{ \begin{aligned} &[\theta_1 + \theta_5(\beta_0 + \beta_2 \gamma_0 + \beta_2 \gamma_c c + \beta_c c) + \theta_4(\gamma_0 + \gamma_c c)] \\ &+ [\theta_4 \gamma_1 + \theta_5(\beta_1 + \beta_3 \gamma_0 + \beta_3 \gamma_c c + \beta_2 \gamma_1)]a_0 \\ &+ \theta_5 \beta_3 \gamma_1 a_0^2 \end{aligned} \right\} (a_1 - a_0) \\
rPSE_{A \rightarrow M_1 \rightarrow Y} &= Q(a_1, a_1, a_0, a_0) - Q(a_1, a_0, a_0, a_0) \\
&= \left\{ \begin{aligned} &(\theta_2 \gamma_1 + \theta_6 \gamma_1(\beta_0 + \beta_2 \gamma_0 + \beta_2 \gamma_c c + \beta_c c)) \\ &+ (\theta_4 \gamma_1)a_1 \\ &+ \theta_6 \gamma_1(\beta_1 + \beta_3 \gamma_0 + \beta_3 \gamma_c c)a_0 \\ &+ \theta_6 \beta_3 \gamma_1^2 a_0^2 \end{aligned} \right\} (a_1 - a_0) \\
rPSE_{A \rightarrow M_2 \rightarrow Y} &= Q(a_1, a_1, a_1, a_0) - Q(a_1, a_1, a_0, a_0)
\end{aligned}$$

$$\begin{aligned}
&= \left\{ \begin{aligned} &(\theta_3 + \theta_6\gamma_0 + \theta_6\gamma_c c)(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c) \\ &+ (\theta_5 + \theta_6\gamma_1)(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c)a_1 \\ &+ (\theta_3 + \theta_6\gamma_0 + \theta_6\gamma_c c)\beta_3\gamma_1 a_0 \\ &+ (\theta_5\beta_3\gamma_1 + \theta_6\beta_3\gamma_1^2)a_1 a_0 \end{aligned} \right\} (a_1 - a_0) \\
rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y} &= Q(a_1, a_1, a_1, a_1) - Q(a_1, a_1, a_1, a_0) \\
&= \left\{ \begin{aligned} &(\theta_3 + \theta_6\gamma_0 + \theta_6\gamma_c c)\beta_2\gamma_1 \\ &+ [(\theta_5\beta_2\gamma_1 + \theta_6\beta_2\gamma_1^2 + (\theta_3 + \theta_6\gamma_0 + \theta_6\gamma_c c)\beta_3\gamma_1)]a_1 \\ &+ (\theta_5\beta_3\gamma_1 + \theta_6\beta_3\gamma_1^2)a_1^2 \end{aligned} \right\} (a_1 - a_0)
\end{aligned}$$

Without the mediator-mediator interaction, i.e. $\theta_6 = 0$, all formulae reduced to

$$\begin{aligned}
&Q(a, a', a'', a''') \\
&= [\theta_0 + \theta_c c + \theta_2(\gamma_0 + \gamma_c c) + \theta_3(\beta_0 + \beta_2\gamma_0 + \beta_2\gamma_c c + \beta_c c)] \\
&+ [\theta_1 + \theta_5(\beta_0 + \beta_2\gamma_0 + \beta_2\gamma_c c + \beta_c c) + \theta_4(\gamma_0 + \gamma_c c)]a \\
&+ \theta_2\gamma_1 a' + \theta_3(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c)a'' + \theta_3\beta_2\gamma_1 a''' \\
&+ \theta_4\gamma_1 aa' + \theta_5(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c)aa'' + \theta_5\beta_2\gamma_1 aa''' + \theta_3\beta_3\gamma_1 a''a''' + \theta_5\beta_3\gamma_1 aa''a''' \\
rPSE_{A \rightarrow Y} &= \left\{ \begin{aligned} &[\theta_1 + \theta_5(\beta_0 + \beta_2\gamma_0 + \beta_2\gamma_c c + \beta_c c) + \theta_4(\gamma_0 + \gamma_c c)] \\ &+ [\theta_4\gamma_1 + \theta_5(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c + \beta_2\gamma_1)]a_0 + \theta_5\beta_3\gamma_1 a_0^2 \end{aligned} \right\} (a_1 - a_0) \\
rPSE_{A \rightarrow M_1 \rightarrow Y} &= \{\theta_2\gamma_1 + \theta_4\gamma_1 a_1\} (a_1 - a_0) \\
rPSE_{A \rightarrow M_2 \rightarrow Y} &= \left\{ \begin{aligned} &\theta_3(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c) + \theta_5(\beta_1 + \beta_3\gamma_0 + \beta_3\gamma_c c)a_1 \\ &+ \theta_3\beta_3\gamma_1 a_0 + \theta_5\beta_3\gamma_1 a_1 a_0 \end{aligned} \right\} (a_1 - a_0) \\
rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y} &= \{\theta_3\beta_2\gamma_1 + [(\theta_5\beta_2\gamma_1 + \theta_3\beta_3\gamma_1)]a_1 + \theta_5\beta_3\gamma_1 a_1^2\} (a_1 - a_0)
\end{aligned}$$

In addition, when interaction terms in outcome model are always zero, i.e. $\theta_4 = \theta_5 = 0$,

then

$$\begin{aligned}
rPSE_{A \rightarrow Y} &= \theta_1(a_1 - a_0); \quad rPSE_{A \rightarrow M_1 \rightarrow Y} = \theta_2\gamma_1(a_1 - a_0); \\
rPSE_{A \rightarrow M_2 \rightarrow Y} &= \{\theta_3\beta_1 + \theta_3\beta_3\gamma_0 + \theta_3\beta_3\gamma_c c + \theta_3\beta_3\gamma_1 a_0\} (a_1 - a_0); \\
\text{and } rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y} &= \{\theta_3\beta_2\gamma_1 + \theta_3\beta_3\gamma_1 a_1\} (a_1 - a_0).
\end{aligned}$$

Furthermore, when interaction terms in the second mediator model is zero, i.e. $\beta_3 = 0$,

then

$rPSE_{A \rightarrow Y} = \theta_1(a_1 - a_0)$; $rPSE_{A \rightarrow M_1 \rightarrow Y} = \theta_2\gamma_1(a_1 - a_0)$; $rPSE_{A \rightarrow M_2 \rightarrow Y} = \theta_3\beta_1(a_1 - a_0)$;
 and $rPSE_{A \rightarrow M_1 \rightarrow M_2 \rightarrow Y} = \theta_3\beta_2\gamma_1(a_1 - a_0)$, which have the same form of the path analysis.

Appendix 3.4. SAS code for randomly interventional analogues of path-specific effects

This SAS code is developed to estimate the randomly interventional analogues of path-specific effects (PSE) of exposure A on a continuous outcome Y with continuous mediators M1 and M2 under the regression models in the “A regression based approach and illustration” section. Suppose we have a dataset named mydata with outcome variable “Y”, exposure variables “A”, two mediators “M1” and “M2”, and three covariates “c1”, “c2”, and “c3”. If there were more or fewer covariates the user would have to modify the second to seven, thirteen, and fourteen lines of the code below to include these covariates. In the fourth line of code, the user has to specify the two levels of A (a1=1 and a0=0) that are being compared and the value of the covariates C at which the effects are to be calculated (cc1=1; cc2=50; cc3=0;), according to the values in the application of interest. The output demonstrates estimates, confidence intervals, and p-value for the total effect, four rPSEs, the effect via the first mediator (with/without the second one), the effect via the second mediator (with/without the first one), the total indirect effect, and the proportion divided by total effect.

```
proc nlmixed data=mydata; /* specify dataset named “mydata” */
parms t0=0 t1=0 t2=0 t3=0 t4=0 t5=0 tc1=0 tc2=0 tc3=0 b0=0 b1=0 b2=0 b3=0
bc1=0 bc2=0 bc3=0 r0=0 r1=0 rc1=0 rc2=0 rc3=0 ss_m1=1 ss_m2=1 ss_y=1; /* para-
meter to be estimated*/

a1=1; a0=0; cc1=1; cc2=50; cc3=0; /*parameter to be intervened*/

/* regression model for mean of all variables */
mu_y=t0 + t1*A + t2*M1 + t3*M2 + t4*A*M1 + t5*A*M2 + tc1*C1 + tc2*C2 +
tc3*C3;

mu_m2 =b0 + b1*A + b2*M1 + b3*A*M1 + bc1*C1 + bc2*C2 + bc3*C3;

mu_m1 =r0 + r1*A + rc1*C1 + rc2*C2 + rc3*C3;

/* score function for all variables*/
```

```

ll_y= -((y-mu_y)**2)/(2*ss_y)-0.5*log(ss_y);
ll_m2= -((m2-mu_m2)**2)/(2*ss_m2)-0.5*log(ss_m2);
ll_m1= -((m1-mu_m1)**2)/(2*ss_m1)-0.5*log(ss_m1);
ll_o= ll_m1 + ll_m2 + ll_y;
model Y ~general(ll_o); /* estimate parameters */
/* calculate all estimate we want */
bcc = bc1*cc1 + bc2*cc2 + bc3*cc3;
rcc = rc1*cc1 + rc2*cc2 + rc3*cc3;
pse0 = ((t1+t5*(b0+b2*r0+b2*rcc+bcc)+t4*(r0+rcc))
        +(t4*r1+t5*(b1+b3*r0+b3*rcc+b2*r1))*a0+t5*b3*r1*a0*a0)*(a1-a0);
pse1 =(t2*r1+t4*r1*a1)*(a1-a0);
pse2 =(t3*(b1+b3*r0+b3*rcc)
        +t5*(b1+b3*r0+b3*rcc)*a1+t3*b3*r1*a0+t5*b3*r1*a1*a0)*(a1-a0);
pse12 =(t3*b2*r1+((t5*b2*r1+t3*b3*r1))*a1+t5*b3*r1*a1*a1)*(a1-a0);
ie1=pse1+pse12;
ie2=pse2+pse12;
ie = pse1+pse2+pse12;
te = pse0+ie;
estimate 'Direct Effect' pse0;
estimate 'Path Specific Effect via M1 alone' pse1;
estimate 'Path Specific Effect via M2 alone' pse2;
estimate 'Path Specific Effect via both M1 and M2' pse12;
estimate 'Path Specific Effect via M1 (with/out M2)' ie1;
estimate 'Path Specific Effect via M2 (with/out M1)' ie2;
estimate 'Total Indirect Effect' ie;
estimate 'Total Effect' te;
estimate 'Proportion Direct Effect' pse0/te;

```

```
estimate 'Proportion via M1' pse1/te;  
estimate 'Proportion via M2' pse2/te;  
estimate 'Proportion via both M1 and M2' pse12/te;  
estimate 'Total Proportion via M1' ie1/te;  
estimate 'Total Proportion via M2' ie2/te;  
estimate 'Proportion via M1 or M2' ie/te;  
run;
```